Learning Linguistic Disjunction

Masoud Jasbi[1], Akshay Jaggi[2], & Michael C. Frank[2]

[1] Harvard University

[2] Stanford University

Author Note

Add complete departmental affiliations for each author here. Each new line herein must be indented, like this line.

Enter author note here.

Correspondence concerning this article should be addressed to Masoud Jasbi, Postal address. E-mail: masoud_jasbi@fas.harvard.edu

Abstract

Research on word learning has discovered constraints, cues, and mechanisms that can help a language learner create successful word-meaning mappings. So far, the literature has mainly focused on the acqusition of content words such as nominals and verbs, leaving functional elements largely understudied. The current study fills this gap by investigating the constraints, cues, and mechanisms that can aid the acqusition of disjunction. Based on naturalistic recordings of parent-child interactions, we argue that children may learn to interpret a disjunction by partition their form-meaning mappings based on salient cues that accompany it in child-directed speech. We first show that children start producing *or* between 18-30 months and by 42 months their productions plateau at a constant rate. We also find that the most likely interpretation of *or* in child-directed speech is exclusive disjunction. However, exclusive interpretations correlated with a rise-fall intonation, and logically inconsistent propositions. In the absence of these two cues, *or* was commonly not exclusive. Our computational modeling shows that a hypothetical learner can successfully interpret an English disjunction by mapping forms to meanings after partitioning the input using the set of salient cues (cue-based) in the context of the utterance (context-depenent). We discuss the implications of our work for current theories of word learning.

Keywords: keywords

Word count: X

Learning Linguistic Disjunction

Introduction

Word learning is the process of detecting a word form, hypothesizing candidate meanings, and mapping the word to its correct meaning (Clark, 1993). As Quine (1960) explained, this process is theoretically complex. Consider a child seeing her father pointing to a fish tank and saying: "*mahi*"! As you probably noticed, *mahi* can mean many things in this context. From "water", "fish", "tail", "smelly", and "look" to "no touching", "few bubbles", and even more complex concepts like "fish-during-the-day" or "fish-water-bubble". Quine argued that the meanings of linguistic utterances are always underdetermined by the behavioral data available to the learner. This problem is known as "the mapping problem", "the gavagai problem", or "indeterminacy of reference".

There are three general ways of tackling this problem. First, word learning may place a priori **constraints** or biases on the hypothesis space (Markman, 1990; Markman & Hutchinson, 1984; Markman & Wachtel, 1988). For example, children may initially assume that words refer to the whole object rather than its parts. Such a constraint would avoid mapping *mahi* to concepts like "tail" or "eye". In addition, children may assume that new words extend to taxonimically related objects and not thematically related ones, therefore avoiding mapping to a concept like "fish-water-bubble". Finally, if the child already knows the word for "water", they may assume that *mahi* cannot be the word for the same concept.

Second, linguistic or sociopragmatic **cues** can add bias for or against some hypotheses. Suppose instead of just saying "*mahi*", the father said "*mahi ro bebin!*". If the child knows "_____ *+ ro + bebin*" as a noun + article + verb combination, then she can analyze *mahi* as a noun and limit its candidate meanings to nominal concepts. If she also knows that "*bebin*" means "look at", she can further limit the candidate meanings to things she can actually look at in the scene. Therefore, prior syntactic or semantic knowledge can cue the meaning of an unknown word (Brown, 1957; Gleitman, 1990). Prior knowledge of communicative acts or human social interaction can also inform word learning. For example, the father's pointing to a fish or looking at it while saying *mahi* can also inform the child of what needs to be attended to for understanding the utterance (Baldwin, 1993; Clark, 2009; Tomasello, 2003).

Third, while each learning instance of a word in isolation may be compatible with innumerable candidate meanings, taken together and aggregating **across situations**, a learner may be able to reduce the indeterminacy substantially (Siskind, 1996; Smith, Smith, & Blythe, 2011; Yu & Smith, 2007). For example, the father may later point to the picture of a fish in a story book and say "*mahi*!" This time there is no "tank" or "bubbles", and the reading may be happening during the night. Therefore, the new observation makes hypotheses like "tank", "few bubbles", or "fish-during-the-day" less plausible. If children track which hypotheses fair best across several naming instances, they have a better chance of narrowing down the hypothesis space to the correct meaning.

Constraints, cues, and cross-situational learning can also operate in conjunction in language acquisition (Hollich et al., 2000). Furthermore, the role and prominence of each factor may vary for different classes of words. So far, nominals have received most of the attention in research on cues and constraints that aid word learning. Function words, on the other hand, have remained largely understudied. In this study, we focus on the role of constraints, cues, and cross-situational word learning on the acquistiion of the disjunction word *or*. In the next two parts of this section, we first summarize previous work on the acquisition of disjunction, and second, summarize our account based on the data presented in the current study.

Previous Studies

To our knowledge, only one study has looked at spontaneous productions of *and* and *or* in parents' and children's speech. Morris (2008) investigated children between the ages of 2;0 and 5;0, using 240 transcriptions of audiotaped exchanges obtained from the CHILDES database. Each connective was analyzed with respect to its frequency, sentence type, and meaning (or use). The study found that overall, *and* was approximately 12.8 times more likely to be produced than *or*. The connective *and* appeared predominantly in statements (more than 90% of the time) while *or* was most common in questions (more than 85% of the time). Children started producing *and* at 2 years and *or* at 2.5 years of age.

Regarding the meaning of the connectives, Morris (2008) adopted a usage-based (item-based) approach (Levy & Nelson, 1994; Tomasello, 2003) and predicted that children start producing connectives with a single "core meaning" (also referred to as "use" or "communicative function"). He predicted that the core meaning mirrors the most frequent meaning of the connective in child-directed speech. Children acquire the less frequent meanings of the connectives as they grow older. He found that children started producing *and* as conjunction at 2, and *or* as exclusive disjunction at 2.5 years of age. In line with the predictions of the usage-based account, he found that these two meanings are the most frequent meanings in parents' speech. For disjunction, 75-80% of the *or*-examples children heard recevied an exclusive interpretation. Finally, as children grew older, they started using connectives to convey additional meanings such as inclusive disjunction for *or* and temporal conjunction for *and*. Overall in adult speech, the inclusive use of *or* was extremely rare, and children barely produced it even at age 5. Morris (2008) argued that the development of connectives conforms to the predictions of a usage-based account and that in the first five years of children's development, the (core) meaning of disjunction is exclusive.

However, a series of experimental studies have found that preschool children are more likely to interpret *or* as inclusive in a variety of linguistic contexts such as negative sentences (Crain, Gualmini, & Meroni, 2000), conditional sentences (Gualmini, Crain, & Meroni, 2000), restriction and nuclear scope of the universal quantifier *every* (Chierchia, Crain, Guasti, Gualmini, & Meroni, 2001; Chierchia et al., 2004), nuclear scope of the negative quantifier *none* (Gualmini & Crain, 2002), restriction and nuclear scope of *not every* (Notley et al., 2012a), and prepositional phrases headed by *before* (Notley et al., 2012b). These studies almost unanimously claim that at least in declarative sentences, the inclusive interpretation of *or* emerges earlier than the exclusive interpretation.

The findings of these studies and Morris (2008) give rise to a paradox: how can children learn to interpret linguistic disjunction as inclusive, if they rarely hear it as inclusive? One way to addresses this paradox is logical nativism (Crain, 2012; Crain & Khlentzos, 2008, 2010). It proposes that the language faculty constrains the connective meanings entertained by the learner to those used in classical logic: negation, conjunction, and inclusive disjunction. Crain (2012) considered it unlikely that children learn the meaning of *or* from the examples they hear in adult usage. Instead, he argued that children rely on an innate knowledge that the meaning of disjunction words in natural languages must be inclusive. In other words, upon hearing a connective word, children consider inclusive disjunction as a viable candidate for its meaning but not exclusive disjunction. In this account, the exclusive interpretation emerges as part of children's pragmatic development after they have mastered the inclusive semantics of disjunction.

While logical nativism addresses the paradox of learning disjunction, it does not provide an explanation for cases where children interpret disjunction as exclusive. Morris (2008) reported that in his study, the vast majority of children used *or* in its exclusive sense. This is not expected if preschool children consider disjunction to be inclusive. Second, other experimental studies, especially those testing disjunction in commands, find that preschool children interpret it as exclusive (Braine & Rumain, 1981; Johansson & Sjolin, 1975). For example, in response to a command such as "give me the doll or the dog", children as young as three- and four-years-old give one of the objects and not both. In its current version, the nativist account does not explain such cases.

Figure 1 summarizes the usage-based and nativist approaches to the acquisition of disjunction. The major difference between them is their assumptions on the learners' semantic hypothesis space for *or*. The usage-based account considers a wide array of meanings to be available for mapping, including different flavors of conjunction such as "temporal conjunction" (e.g. Bob pressed the key and (then) the door opened) and "explanatory conjunction". The nativist account places more constraints ont the hypothesis space and limits it to binary logical connectives of standard propositional logic: inclusive disjunction, conjunction, and material implication. Both accounts agree that the input favors the exclusive interpretation of disjunction. The usage-based account concludes that children's early mappings mirror this input. The nativist account suggests that innate biases towards the inclusive meaning and against the exclusive interpretation result in an inclusive semantics for *or* in children's early mappings.

Current Study

In this study, we provide an alternative solution to the paradox of learning disjunction. The main claim of this paper is that child-directed speech contains salient cues that accompany a linguistic disjunction and can help a learner successfully interpret it – for example as exclusive or inclusive. We support this hypothesis using three studies. Study 1 does not directly support our main claim but provides the necessary basic information. It investigates the distribution of *and* and *or* in parent-child interactions to address the following basic questions: how often do children hear or produce *or*? and when do they start producing it? Using a large corpus of parent-child interactions, we found that children hear 1-2 examples of *or* in every thousand words parents produce. They start producing it

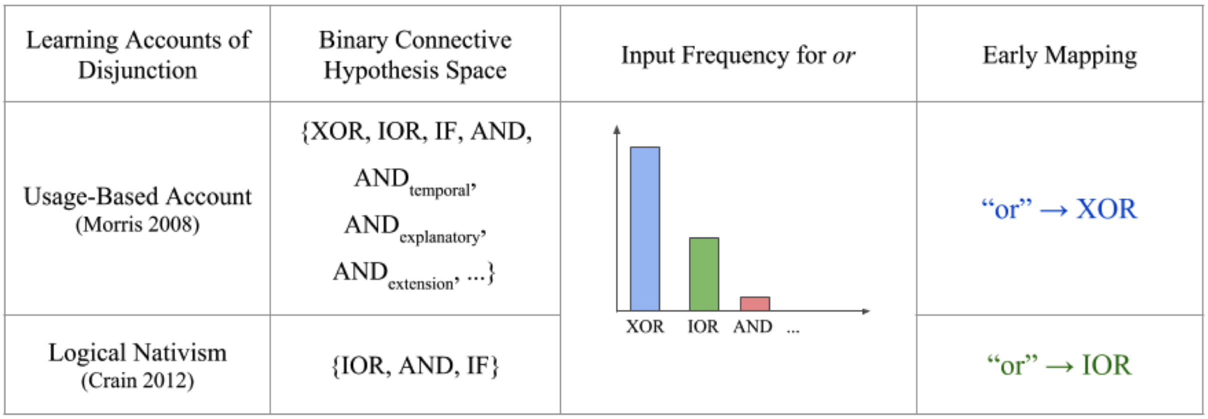

Figure 1. Summary of the usage-based and nativist approaches to the acquisition of disjunction.

themselves between 18-30 months, and by 42 months they reach a rate of one *or* per thousand words. Studies 2 and 3 provide support for the two parts of our main claim: presence of cues, and their utility in learning. In study 2, we ask: what interpretations can *or* have in child-directed speech? We annotated examples of *or* and found that its most likely interpretation is exclusive disjunction, as Morris (2008) had concluded. However, we also found that exclusive interpretations correlated strongly with two cues: rise-fall prosody, and logically inconsistent propositions connected by *or*. In the absence of these cues, *or* was most likely non-exclusive. In our third study, we asked if it is possible to learn the interpreations of *or* from these cues. Using the annotation data of study 2 and a supervised learning task, we showed that a decision-tree classifer can use prosody and consistency of propositions to predict its interpretation with high accuracy.

Based on the results of our studies, we propose a new account for children's acquisition of disjunction. Figure 2 shows the summary of this account which we call "cue-based context-dependent mapping" of disjunction. It is inspired by the usage-based and nativist accounts of disjunction and shares many of their insights. Similar to the nativist account, we assume that the semantic hypothesis space includes binary logical relations. However, we do not limit the hypothesis space further and do not bias the learning towards the inclusive meaning. We will show that the linguistic input can achieve this. Similar to usage based proposals, our account relies on the structure of the input to distinguish between exclusive and inclusive uses of disjunction. We also map more complex constructions to meanings rather than the word *or* directly. The learner can later extract commonalities across mappings to complex forms and extract a core semantics for a particular word like *or*. The major point of departure from previous accounts is the mechanism of learning. While in pervious accounts the most frequent meaning in the input was mapped to the connective word directly, in our account the input is partitioned or broken down by a set of salient cues that designate the context of use. Mapping is done based on the cues that accompany the connective word.

Study 1: Production of "or" in parent-child interactions

In our first study, we looked at the frequencies of *and* and *or* in a corpus of parent-child interactions (CHILDES) with 14,159,609 words. This is a considerably larger corpus than previously used.

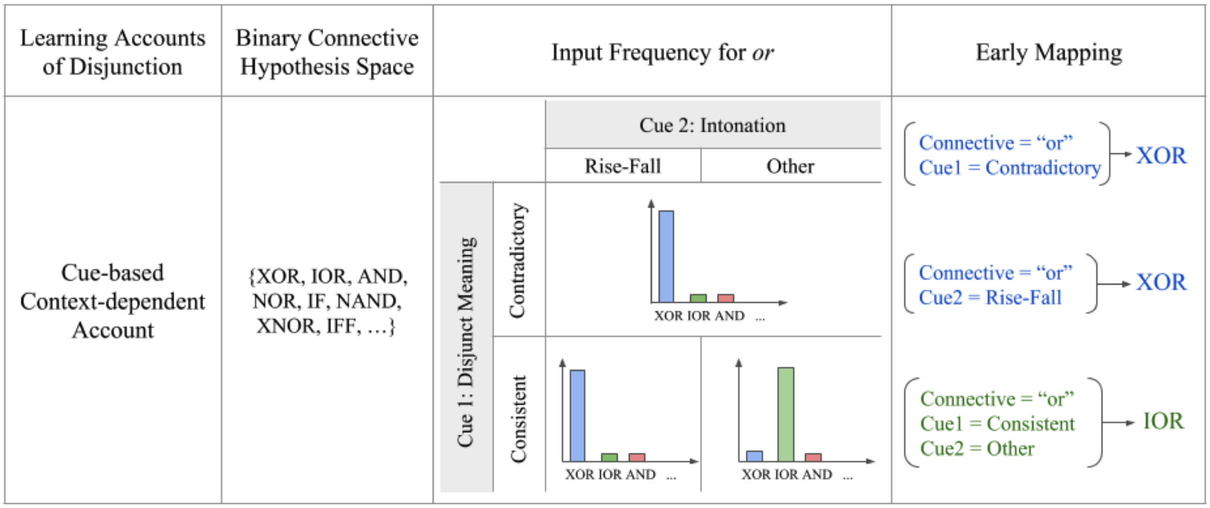

Figure 2. Summary of the usage-based and nativist approaches to the acquisition of disjunction.

Methods

For samples of parents' and children's speech, we used the online database childes-db and its associated R programming package `childesr` (Sanchez et al., 2018). Childes-db is an online interface to the child language components of TalkBank, namely CHILDES (MacWhinney, 2000) and PhonBank. Two collections of corpora were selected: English-North America and English-UK. All word tokens were tagged for the following information: 1. The speaker role (mother, father, child), 2. the age of the child when the word was produced, 3. the type of the utterance the word appeared in (declarative, question, imperative, other), and 4. whether the word was *and*, *or*, or neither.

Exclusion Criteria. First, tokens were coded as unintelligible were excluded (N = 290,119). Second, tokens that had missing information on children's age were excluded (N = 1,042,478). Third, tokens outside the age range of 1 to 6 years were excluded (N = 686,870). We were interested in the 1 to 6 years old age range and there was not much data outside this age range. The collection contained the speech of 504 children and their parents after the exclusions.

Procedure. Each token was marked for the utterance type that the token appeared in. This study grouped utterance types into four main categories: "declarative", "question", "imperative", and "other". Utterance type categorization followed the convention used in the TalkBank manual. The utterance types are similar to sentence types (declarative, interrogative, imperative) with one exception: the category "question" consists of interrogatives as well as rising declaratives (i.e. declaratives with rising question intonation). In the transcripts, declaratives are marked with a period, questions with a question mark, and imperatives with an exclamation mark. It is important to note that the manual also provides terminators for special-type utterances. Among the special type utterances, this study included the following in the category "questions": trailing off of a question, question with exclamation, interruption of a question, and self-interrupted question. The category imperatives also included "emphatic imperatives". The rest of the special type utterances such as "interruptions" and "trailing off" were included in the category "other".

Results

Overall, *and* was about 10 times more likely to occur in parents' speech than *or*. More specifically, *and* occurred 15 times and *or* only 1.5 times per 1000 words. Children produced *and* at the same rate as their parents but produced *or* at a considerably lower rate, only 0.5 per thousand (Figure 3, Left). The developmental trend showed that between 12 to 72 months, production of *and* in parents' speech varied between 10 to 20 per thousand words

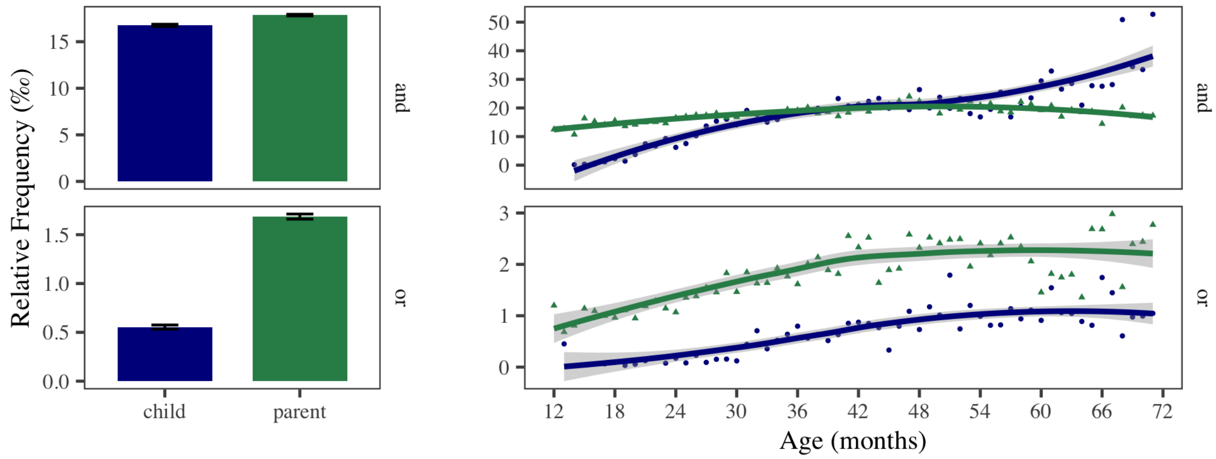

Figure 3. Left: The relative frequency of *and/or* (per mille) in the speech of parents and children. 95% binomial proportion confidence intervals calculated using Agresti-Coull's approximate method. Right: The monthly relative frequency of *and/or* in parents and children's speech between 12 and 72 months (1-6 years).

(Figure 3, Right). Children started producing *and* between 12 and 18 months, and showed a sharp increase in their production until they reached the parent level between 30 to 36 months of age. Their productions stayed close to the parents' production level between 36 and 72 months, possibly surpassing them at 60 months – although due to the small amount of data after 60 months we should be cautious with our interpretation of the trend there. The production of *or* for parents was 1 to 2 per thousand words. Children started producing *or* between 18 to 30 months, steadily increasing their productions until they got close to 1 *or* per thousand words at 48 months (4 years). Their productions plateaued and stayed at this rate until 72 months (6 years).

Children's productions of *or* was different from their production of *and* and parents' production of *or*. Children started producing *or* around 6 months later than they started with *and*. Second, while children's *and* productions showed a steep rise over a year and reached the parent level around 30 months, their *or* productions rose slowly and did not reach the parent level even at 6 years of age. What factors cause these differences? We consider three possibilities here: frequency, conceptual complexity, and usage.

First, *and* is a far more frequent connective than *or*. Goodman, Dale, and Li (2008) argue that within the same syntactic category, words with higher frequency in child-directed speech are acquired earlier. The conjunction word *and* is at least 10 times more likely to occur than *or* so earlier acquisition of *and* is consistent with the effect of frequency on age of acquisition. Second, research on concept attainment has suggested that the concept of conjunction is easier to conjure and possibly acquire than the concept of disjunction. In experiments that participants are asked to detect the pattern of classification in some cards, they can detect a conjunctive classification faster than a disjunctive one (Neisser & Weene, 1962). Therefore, it is possible that children discover the concept that corresponds to the meaning of *and* faster and start to produce it earlier, but they need more time to attain the concept corresponding to the meaning of *or*.

A third possibility is that the developmental difference between *and* and *or* is at least partly due to their different usages. Parent-child interactions are not symmetrical and what parents would like to communicate to children is different from what children would like to communicate to parents. This asymmetry can result in different distribution of speech acts between parents and children and consequently functional elements that constitute them. Here we present evidence that suggests *or* is affected in this way.

First, we found that *or* was more likely to occur in questions than in declaratives (Figure 4, Left). This is in contrast to *and* which was more likely to occur in declaratives. Second, parents asked more questions from children than children did from parents, and children produced more declaratives than parents (Figure 4, Right). In fact, questions had their own developmental trajectory, emerging in the second year of children's lives and reaching a relatively constant rate of about 15% of children's utterances in their fourth year. However, parents produce a constant rate of questions which is about 25% of their utterances. Therefore, parent-child interaction provides more opportunities for parents to ask

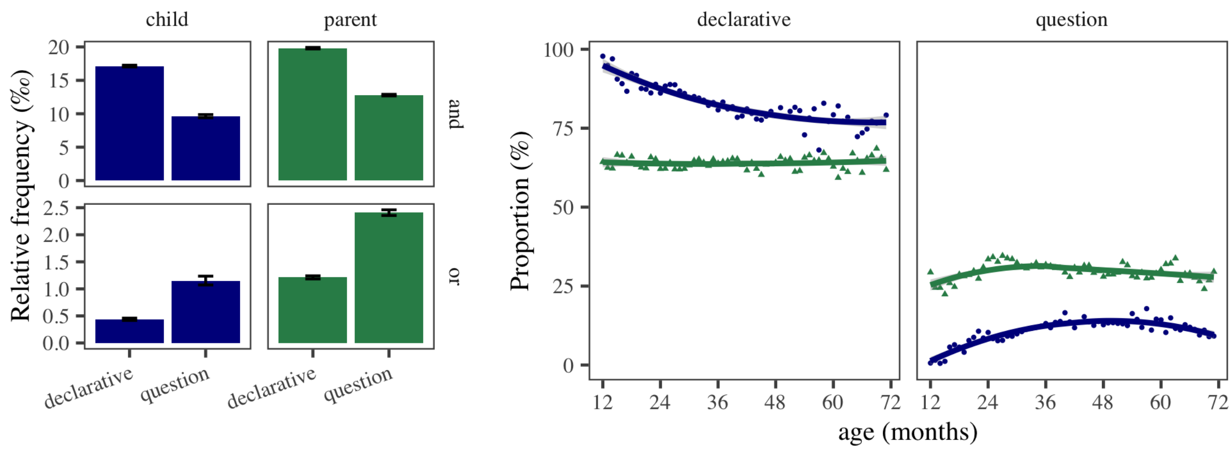

Figure 4. Left: Relative frequency of *and/or* (per mille) in declaratives, imperatives, and interrogatives for parents and children. Right: Percentage of declaratives to questions in parent-child interactions by age.

questions and produce *or*, than children.

Figure 5 shows the developmental trends for the relative frequencies of *and* and *or* in questions and declaratives. Comparing *and* in declaratives and questions, we see that the onset of *and* productions were slightly delayed for questions. But in both declaratives and questions, *and* productions reached the parent level around 30 months (2.5 years). For *or*, we see a similar delay in questions compared to declaratives. Children started producing *or* in declaratives at around 18 months but they started producing *or* in questions at 24 months. Production of *or* increased in both declaratives and questions until it reached a constant rate in declaratives between 48 and 72 months. The relative frequency of *or* in questions continued to rise until 60 months. Comparing Figure 3 and 5, children were closer to the adult rate of production in declaratives than questions.

To test these observations more formally, we used a linear regression model with the relative frequency of *or* as the dependent variable and children's age, speaker (child vs. parent), utterance type (declarative vs. question), and their interactions as predictors.

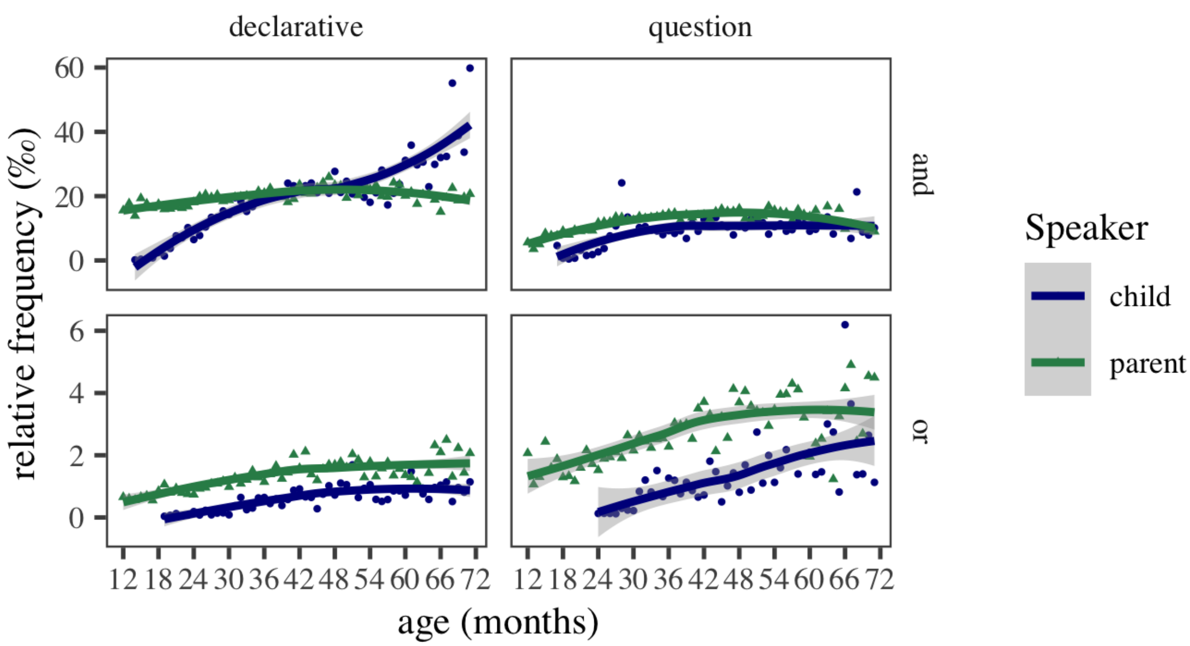

Figure 5. Relative frequency of *and/or* in declaratives and questions for parents and childern between the child-age of 12 and 72 months (1-6 years).

The intercept was set to children's productions in declaratives. Table 1 presents the coefficient estimates of the model. Overall, the model suggests that parents and children produced more *or* as children grew older and parents produced more instances of *or* than children. However, the increase in production of *or* was more steep in questions. The largest significant effect was the interaction of speaker and utterance type. Parents produced disjunctions more frequently in quesions than in declaratives. These results are consistent with the hypothesis that frequency and distribution of *or* is partly affected by the development of questions in parent-child interactions.

Conclusion

In a large-scale quantitative analysis of parents and children's productions of *and* and *or*, we found that children started producing *and* in the second year of their lives, and

Table 1

Estimated cofficients for the linear model with children's age, speaker (child vs. parent), utterance type (declarative vs. question), and their interactions as predictors. Relative frequency of disjunction produciton was the dependent variable.

Coefficients	Estimate	Std. Error	t value	Pr(>$\lvert$t$\rvert$)
age	0.02	0.01	3.54	0.00
question	-0.77	0.39	-1.96	0.05
parent	0.72	0.32	2.24	0.03
age*question	0.03	0.01	3.96	0.00
age*parent	0.00	0.01	0.21	0.83
question*parent	1.40	0.48	2.91	0.00
age*question*parent	-0.01	0.01	-1.30	0.20

quickly reached their parents' rate of production by two and a half. Their production of disjunction was delayed by six months on average: they started producing *or* between 1.5 and 2.5 years of age, and around 3.5 years, they reached a relatively constant rate of production below that of their parents. We considered three possible causes for disjunction's delay and lower rate of production: the higher frequency of *and*, the conceptual and mapping complexity of *or*, and the asymmetry in speech acts produced by parents and children. We provided evidence for the last cause. We showed that parents produced more questions than children, and that *or* was more likelly to occur in questions. Therefore, parents' speech contained more *or* partly due to the fact that parents asked more questions.

Study 2: Interpretations of disjunction in child-directed speech

In this study we selected a subset of connective examples in child-directed speech from study 1 to closely examine the interpretations they recieve. Research in formal semantics has shown that the interpretation of disjunction depends on several factors including prosody (Pruitt & Roelofsen, 2013), logical consistency of the propositions being connected (Geurts, 2006), pragmatic and scalar reasoning (Grice, 1989). Our main claim here is that in child-directed speech, exclusive interpretations of *or* correlate with rise-fall prosody and logically inconsistent propositions. In the absence of these two factors, *or* is most likely "not exclusive".

Methods

This study used the Providence corpus (Demuth, Culbertson, & Alter, 2006) available via the PhonBank section of the TalkBank.org archive. The corpus was chosen because of its relatively dense data on child-directed speech as well as the availability of audio and video recordings that would allow annotators access to the context of the utterance. The corpus was collected between 2002 and 2005 in Providence, Rhode Island. Table 2 in appendix reports the name, age range, and the number of recording sessions for the children in this study. All children were monolingual English speakers and were followed between the ages of 1 and 4 years. Based on Study 2, this is the age range when children develop their early understanding of *and* and *or*. The corpus contains 364 hours of biweekly hour-long interactions between parents and children.

Exclusion Criteria. We excluded data from Ethan since he was diagnosed with Asperger's Syndrome at age 5. We also excluded all examples found in conversations over the phone, adult-adult conversations, and utterances heard from TV or radio. We did not count such utterances as child-directed speech. We excluded proper names and fixed forms such as "Bread and Circus" (name of a local place) or "trick-or-treat" from the set of examples to be annotated. Such forms could be learned and understood with no actual understanding of the connective meaning. We counted multiple instances of *or* and *and* within the same disjunction/conjunction as one instance. The reasoning was that, in a coordinated structure, the additional occurrences of a connective typically did not alter the annotation categories, and most importantly the interpretation of the coordination. For example, there is almost no difference between "cat, dog, and elephant" versus "cat and dog and elephant" in interpretation. In short, we focused on the "coordinated construction" as a unit rather than on every separate instance of *and* and *or*. Instances of multiple connectives in a coordination were rare in the corpus.

Procedure. All utterances containing *and* and *or* were extracted using the CLAN software and automatically tagged for the following: (1) the name of the child; (2) the transcript address; (3) the speaker of the utterance (father, mother, or child); (4) the child's birth date, and (5) the recording date. Since the focus of the study was mainly on disjunction, we annotated instances of *or* in all the child-directed speech from the earliest examples to the latest ones found. Given that the corpus contained more than 10 times the number of *and*'s than *or*'s, we randomly sampled 1000 examples of *and* to match 1000 examples of *or*. Here we report the results on 627 examples of *and* and 608 examples of *or*.

Annotation Categories. Every extracted instance of *and* and *or* was manually annotated for 7 categories: connective interpretation, intonation type, utterance type, syntactic level, conceptual consistency, communicative function, and answer type. We briefly explain how each annotation category was defined. Further details and examples are provided in the appendix section.

1. *Connective Interpretation*

This annotation category was the dependent variable of the study. Annotators listened to coordinations such as "A or B" and "A and B", and decided the intended interpretation of the connective with respect to the truth of A and B. We used the sixteen binary connective meanings shown in Figure 6. Annotators were asked to consider the two propositions raised by the coordinated construction, ignoring the connective and functional elements such as negation. Consider the following sentences containing *or*: "Bob plays soccer or tennis" and "Bob doesn't play soccer or tennis". Both discuss the same two propositions: A. Bob playing

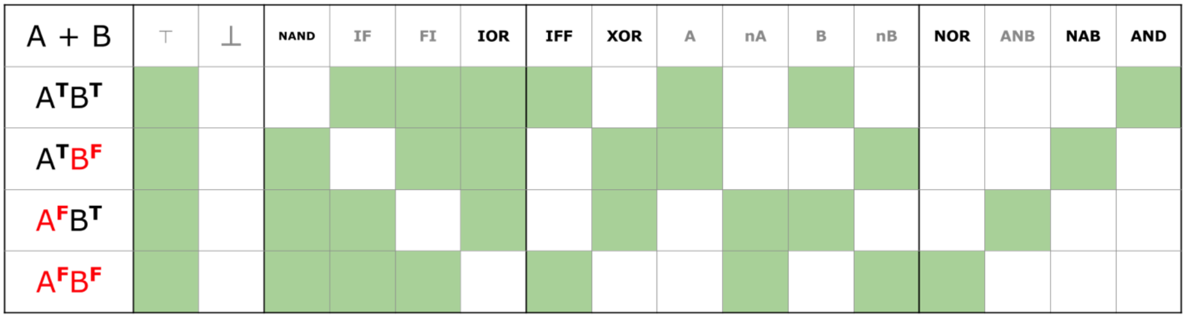

Figure 6. The truth table for the 16 binary logical connectives. The rows represent the set of situations where bot A and B, A, B, or, neither propositions are true. The columns represent the 16 possible connectives and their truth conditions. Green cells represent true situations.

soccer, and B. Bob playing tennis. However, the functional elements combining these two propositions result in different interpretations with respect to the truth of A and B. In “Bob plays soccer or tennis” which contains a disjunction, the interpretation is that Bob plays one or possibly both sports (IOR). In “Bob doesn’t play soccer or tennis” which contains a negation and a disjunction, the interpretation is that Bob plays neither sport (NOR). For connective interpretations, the annotators first reconstructed the coordinated propositions without the connectives or negation and then decided which propositions were implied to be true/false.

2. *Conceptual Consistency*

Propositions stand in complex conceptual relations with each other. For example, have logical, temporal, and causal relation with each other. For conceptual consistency, annotators decided whether the propositions that made up the coordination could be true at the same time or not. If the two propositions could not be true at the same time and resulted in a contradiction, they were marked as inconsistent. Our annotators used the following diagnostic to decide the consistency of the disjuncts: Two disjuncts were marked as inconsistent if replacing the word *or* with *and* produced a contradiction. For example, changing "the ball is in my room *or* your room" to "the ball is in my room *and* your room" produces a contradiction because a ball cannot be in two rooms at the same time.

It is important to discuss two issues regarding conceptual consistency. First, our diagnostic for consistency was quite strict. In many cases, propositions are not inconsistent in this sense but they are implausible. For example, drinking both tea and coffee at the same time is not inconsistent, but is unlikely. It is possible that many exclusive interpretations are based on such judgments of implausability. Second, if the coordinands are inconsistent, this does not necessarily mean that the connective interpretation must be exclusive. For example, in a sentence like "you could stay here or go out", the alternatives "staying here" and "going out" are inconsistent. Yet, the overall interpretation of the connective could be conjunctive: you could stay here AND you could go out. The statement communicates that both possibilities hold. This pattern of interaction between possibility modals like *can* and disjunction words like *or* are often discussed under "free-choice inferences" in the semantics and pragmatics literature (Kamp, 1973; Von Wright, 1968). Another example is unconditionals such as "Ready or not, here I come!". The coordinands are contradictions: one is the negation of the other. However, the overall interpretation of the sentences is that in both cases, the speaker is going to come.

3. *Utterance Type*

Annotators decided whether an utterance was an instance of a declarative, an interrogative, or an imperative. Occasionally, we found examples with different utterance types for each coordinand. For example, a mother could say "put your backpack on and I'll be right back", where the first cooridnand is an imperative and the second a declarative. Such examples were coded for both utterance types with a dash inbetween: imperative-declarative. Table 5 in the appendix provides the detailed definitions and examples for each utterance type.

4. *Intonation Type*

Annotators listened to the utterances and decided whether the intonation contour on the coordination was flat, rise, or rise-fall. Table 4 in the appendix shows the definitions and examples for these intonation types. In order to judge the intonation of the sentence accurately, annotators were asked to construct all three intonation contours for the same sentence and see which one is closer to the actual intonation of the utterance. For example, to judge the sentence “do you want orange juice↑ or apple juice↓?”, they reconstructed the sentence with the prototypical flat, rising, and rise-fall intonations and checked to see which intonation is closer to the actual one.

5. *Syntactic Level*

Annotators marked whether the coordination was at the clausal level or at the sub-clausal level. Clausal level was defined as sentences, clauses, verb phrases, and verbs. Coordination of other categories was coded as sub-clausal. This annotation category was introduced to check the hypothesis that the syntactic category of the coordinands may influence the interpretation of a coordination. For example, a sentence like “He drank tea or coffee” is less likely to be interpreted as exclusive than “He drank tea or he drank coffee.” The clausal vs. sub-clausal distinction was inspired by the fact that in many languages, coordinators that connect sentences and verb phrases are different lexical items than those that connect nominal, adjectival, or prepositional phrases (see Haspelmath, 2007).

6. *Communicative Functions*

We constructed a set of categories that captured particular usages or communicative functions of the words *or* and *and*. They include descriptions, directives, preferences, identifications, definitions-examples, clarifications, repairs, and a few others shown in Table 8 in appendix. These communicative functions were created using the first 100 examples and then they were used for the classification of the rest of the examples. Some communicative functions are general and some are specific to coordination. For example, directives are a general class while conditionals (e.g. Put that out of your mouth, or I'm gonna put it away) are more specific to coordinated constructions. It is also important to note that the list is not unstructured. Some communicative functions are subtypes of others. For example, “identifications” and “unconditionals” are subtypes of “descriptions” while “conditionals” are a subtype of directives. Furthermore, “repairs” seem parallel to other categories in that any type of speech can be repaired. We do not fully explore the details of these functions in this study but such details matter for a general theory of acquisition that makes use of the speaker's communicative intentions as early coarse-grained communicative cues for the acquisition of fine-grained meaning such as function words.

7. *Answer Type*

Whenever a parent's utterance was a polar question, the annotators coded the utterance for the type of response it received from the children. This annotation category was different from others because it was not used as a cue for learning disjunction. Instead, it was used as an opportunity to assess, albeit in a limited and indirect way, the comprehension of children in the same corpus. Table 9 in the appendix shows the answer types in this study and their definitions and examples. Utterances that were not polar questions were simply coded as NA for this category. If children responded to polar questions with “yes” or “no”, the category was YN and if they repeated with one of the coordinands the category was AB. If children said yes/no and followed it with one of the coordinands, the answer type was determined as YN (yes/no). For example, if a child was asked “Do you want orange juice or apple juice?” and the child responded with “yes, apple juice”, our annotators coded the response as YN. The reason is that in almost all cases, if a simple yes/no response is felicitous, then it can also be optionally followed with mentioning a disjunct. However, if yes/no is not a felicitous response, then mentioning one of the alternatives is the only appropriate answer. For example, if someone asks "Do you want to stay here or go out?" a response such as "yes, go out" is infelicitous and a better response is simply "go out". Therefore, we counted responses with both yes/no and mentioning an alternative as a yes/no response.

8. *Negation and Modals*

Finally, a script was used to automatically mark utterances for whether they contain sentential negation (*not/n't*) or any modal auxiliary such as *maybe*, *can*, *could*, *should*, *would*, or *need to*. This allowed us to see how the presence or absence of negation or modals could affect the overall interpretation of the utterance.

Inter-annotator Reliability. To train annotators and confirm their reliability for disjunction examples, two annotators coded the same 240 instances of disjunction. The inter-annotator reliability was calculated over 8 iterations of 30 examples each. After each iteration, annotators met to discuss disagreements and resolve them. They also decided whether the category definitions or annotation criteria needed to be made more precise. Training was completed after three consecutive iterations showed substantial agreement between the annotators for all categories (Cohen's $\kappa > 0.7$). Further details on inter-annotator reliability are presented in the appendix section.

Results

We start with the category "answer type". This category can help us understand if children in the providence corpus provided appropriate answers to questions with disjunction. Figure 7 (Left) shows the monthly proportions of "yes/no" (Y/N) and alternative (AB) answers between the ages of 1 and 3 years. Initially, children provided no answer to questions, but by the age of 3 years, the majority of such questions received a yes/no (YN)

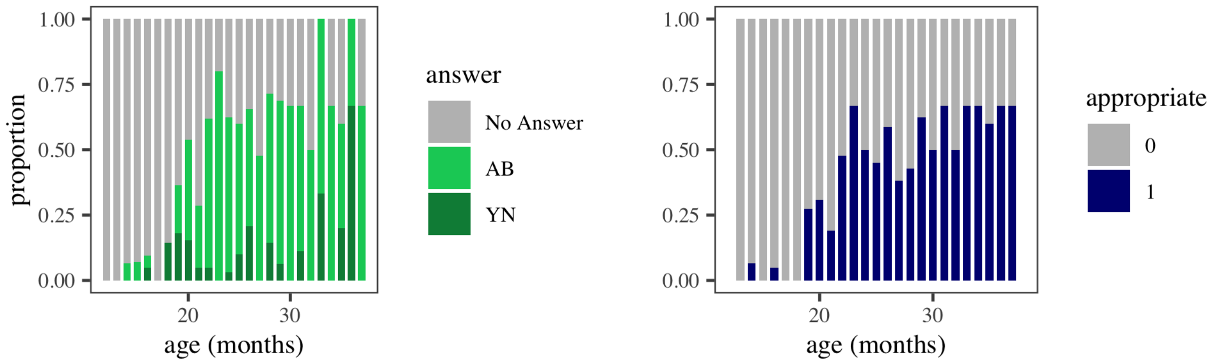

Figure 7. Left: Monthly proportions of children's yes/no (YN) and alternative (AB) answers to questions with *or*. Right: Monthly proportions of children's appropriate answers to questions with *or*.

or alternative (AB) answer. To assess how often these answers were appropriate, we defined appropriate answers the following way: an alternative (AB) answer is appropriate for an alternative question (one with "or" and a rise-fall intonation). A yes/no answer (YN) is appropriate for a yes/no (polar) question (one with *or* and a rising intonation). Of course this classification is strict and misses some nuanced cases, but nevertheless provides a useful conservative estimate. The right side of Figure 7 shows the monthly proportion of children's appropriate answers between the ages of 1 and 3. The results show that even with a conservative measure, children show an increase in the proportion of their appropriate answers to questions containing *or* between 20 to 30 months of age (roughly 2 and 3 years of age). This in turn suggests that initial form-meaning mappings for disjunction is formed in this age range. The rest of this section discusses the cues that can assist children create successful form-meaning mappings.

First, we look at our dependent variable, namely "connective interpretations". Figure 8 (Left) shows the overall distribution of the connective interpretations in child-directed speech regardless of the connective word. The most common interpretation was conjunction (AND, 55%) followed by exclusive disjunction (XOR, 31%). Figure 8 (Right) shows the distribution

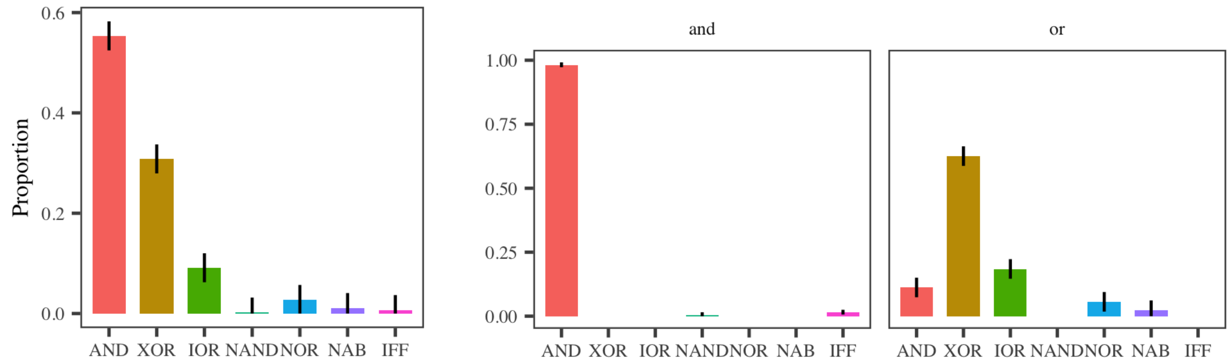

Figure 8. Left: Connective interpretations in child-directed speech. Right: Connective interpretations broken down by lexical items *and* (conjunction) and *or* (disjunction).

of connective interpretations broken down by the connective word used: *and* vs. *or*[1]. Almost all instances of the connective *and*, were interpreted as conjunction (AND). There were also a small number of NAND interpretations (e.g. "don't swing that in the house and hit things with it") and IFF interpretations (e.g. "come here and I'll show you") in our sample. For the connective *or*, the most frequent interpretation was exclusive disjunction (XOR, 62%) followed by inclusive disjunction (IOR, 18%) and conjunction (AND, 11%). There were also a small number of NOR (e.g. "you never say goodbye or thank you") and NAB interpretations (e.g. "those screws, or rather, those nuts"). Overall, these results are consistent with the findings of Morris (2008) who concluded that exclusive disjunction is the most common interpretation of *or*. Therefore, by simply associating the most common interpretations with the connective words, a learner is expected to learn *and* as conjunction, and *or* as exclusive disjunction (Crain, 2012; Morris, 2008).

However, the learning outcome might be different if factors other than the connective word are also considered. In what follows, we investigate how different annotation categories introduced earlier correlate with the interpretations of *or*. We set *and* aside because it was

[1] All the confidence intervals shown in the plots for this section are simultaneous multinomial confidence intervals computed using the Sison and Glaz (1995) method.

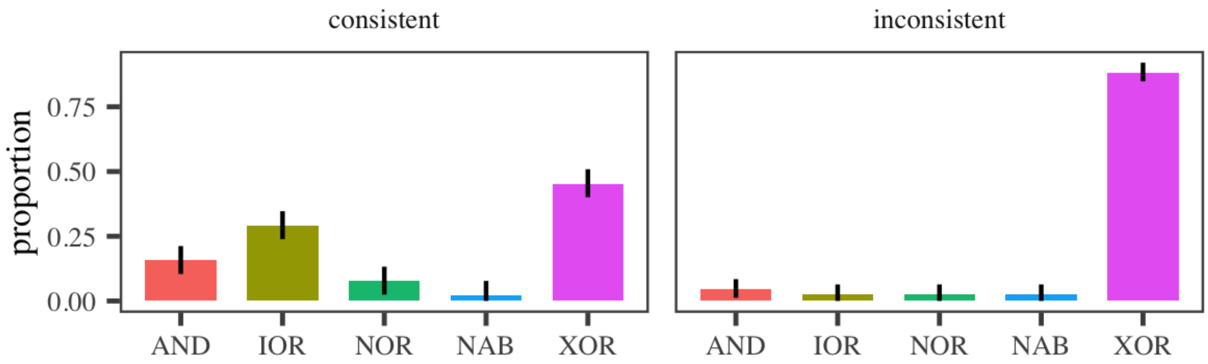

Figure 9. Interpretations of disjunction with consistent vs. inconsistent disjuncts.

almost always interpreted as conjunction (AND). Figure 9 shows the proportions of connective interpretations in disjunctions with consistent vs. inconsistent disjuncts. When the disjuncts were consistent (i.e. could be true at the same time), the interpretation could be exclusive (XOR), inclusive (IOR), or conjunctive (AND). When the disjuncts were inconsistent, a disjunction almost always received an exclusive (XOR) interpretation. This suggests that the exclusive interpretation of a disjunction often stems from the inconsistent or contradictory nature of the disjuncts themselves[2].

Next we focus on cases of disjunction with consistent disjuncts. Figure 10 shows their interpretations in declarative, interrogative, and imperative sentences. Interrogatives selected for exclusive and inclusive interpretations. Imperatives were more likely to be interpreted as inclusive (IOR), but declaratives could receive almost any interpretation: conjunctive (AND), exclusive (XOR), inclusive (IOR), or even that "neither" disjunct was true (NOR). A common example of inclusive imperatives was invitation to action such as "Have some food

[2]It should be noted here that in all *and*-examples, the disjuncts were consistent. This is not surprising given that inconsistent meanings with *and* result in a contradiction. The only exception to this was one example where the mother was mentioning two words as antonyms: "short and tall". This example is quite different from the normal utterances given that it is meta-linguistic and list words rather than asserting the content of the words.

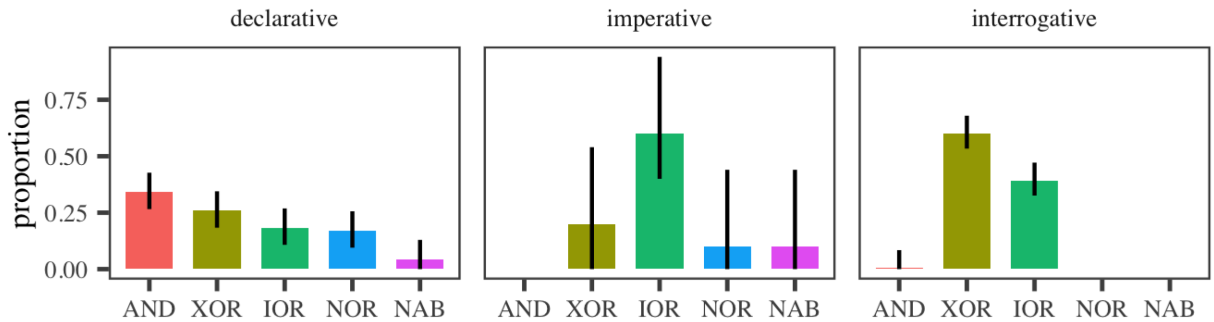

Figure 10. Interpretations of disjunction with consistent disjuncts in interrogative, imperative, and declarative utterances.

or drink!". Such invitational imperatives seem to convey inclusivity (IOR) systematically. They are often used to give the addressee full permission with respect to both alternatives. It can in fact be odd to use them to imply exclusivity (e.g. "Have some food or drink, but not both!"), and they are not conjunctive either, i.e inviting the addressee to do both actions (e.g. "Have some food, and have some drink!").

While interrogatives selected for exclusive and inclusive interpretations, their intonation could distinguish between these two readings. Figure 11 shows the interpretations of consistent disjunction in three intonational contours: flat, rise, and rise-fall. The rise and rise-fall contours are typical of interrogatives. The results show that, a disjunction with a rise-fall intonation is most likely interpreted as exclusive (XOR). If the intonation is rising, a disjunction is most likely inclusive (IOR). Finally, a disjunction with a flat intonation (typical of declaratives and imperatives) could be interpreted as exclusive (XOR), conjunctive (AND), inclusive (IOR), or neither (NOR). These results replicate Pruitt and Roelofsen (2013)'s experimental findings on the role of intonation in the interpretation of polar and alternative questions.

Next we focus on consistent disjunctions with flat intonation. Figure 12 breaks down

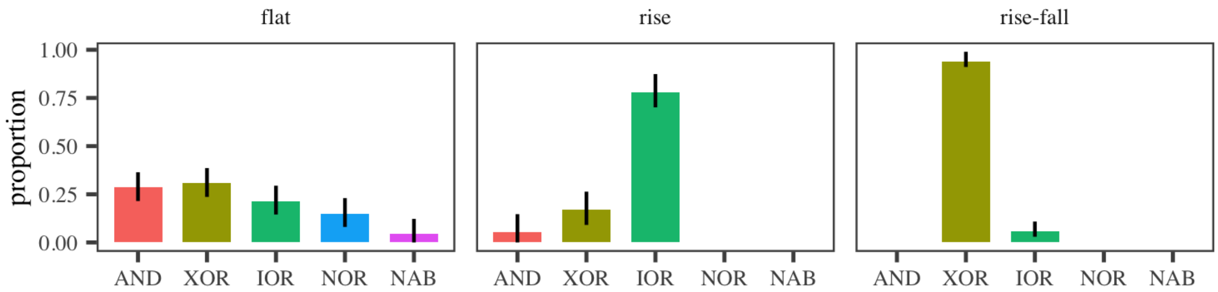

Figure 11. Interpretations of disjunction with consistent disjuncts and flat, rising, or rise-fall intonation.

the interpretations based on whether the utterance contained negation or modals. The results show that in the presence of a modal such as *can* or *maybe*, it was more likely for a disjunction to have a conjunctive interpretation. This is consistent with the literature on free-choice inferences in formal semantics and pragmatics (Kamp, 1973), which shows statements such as "you can have tea or coffee" is interpreted conjunctively as "you can have tea *and* you can have coffee". When the utterance contained a negation, the disjunction could be interpreted as exclusive (XOR) or neither (NOR). These two interpretations correspond to the scope relations between negation and disjunction. If negation scopes above disjunction, we get a neither (NOR) interpretation (e.g. "I do not eat cauliflower, cabbage or baked beans.") But if disjunction scopes above negation, the likely interpretation is exclusive (e.g. don't throw it at the camera or you're going in the house.) These results also suggest that a learner who tracks co-occurences of *or* with negative morphemes can potentially learn about the scope interaction of disjunction and negative particles in their native language.

Finally, we visit the last two remaining categories: syntactic level and communicative functions. For these categories, we show connective interpretations over all instances of disjunction. Figure 13 shows connective interpretations, broken down by syntactic level. The results suggest a possible small effect of clausal level disjuncts. Disjunctions were more likely

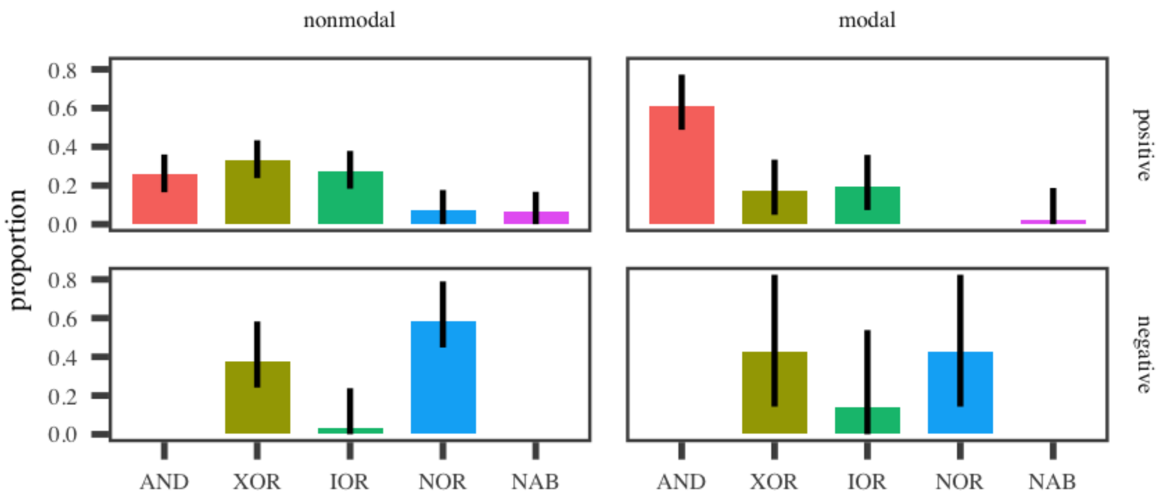

Figure 12. Distribution of connective interpretations for consistent disjuncts with flat intonation.

to be interpreted as exclusive if their disjuncts were clauses or verbs rather than nominals, adjectives, or prepositions (all sub-clausal units). As explained before, the intuition is that a sentences such as "They had tea or coffee" is less likely to be exclusive than "they had tea or they had coffee" However, our understanding is that compared to other factors such as intonation and consistency, the effect of syntactic level was very small. As we shall see in Study 3, a computational learning model did not find syntactic level to be of much use for classifying instances of disjunction as exclusive, above and beyond what other annotation categories offered.

Figure 14 shows connective interpretations in the 10 different communicative functions we defined. The results show that certain functions increase the likelihood of some connective interpretations. An exclusive interpretation (XOR) is common in acts of clarification, identification, stating/asking preferences, stating/asking about a description, or making a conditional statements. These results are consistent with expectations on the communicative intentions that these utterances carry. In clarifications, the speaker needs to know which of

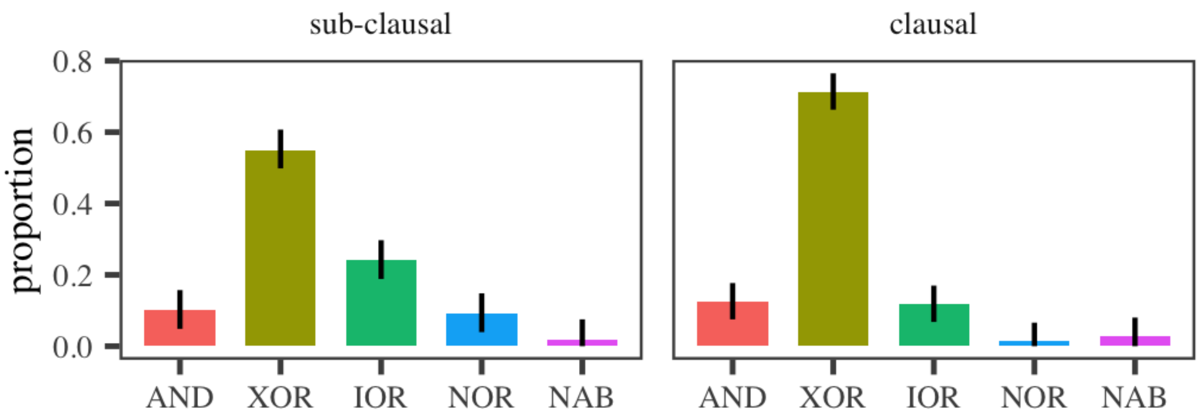

Figure 13. Top: Interpretations of clausal vs. sub-clausal disjunction. Down: Interpretations of clausal vs. sub-clausal disjunction in declaratives with consistent disjuncts.

two alternatives the other party meant. Similarly in identifications, speaker needs to know which category does a referent belongs to. In preferences, parents seek to know which of two alternatives the child wants. Even though descriptions could be either inclusive or exclusive, in the current sample, most descriptions were questions about the state of affairs and required the child to provide one of the alternatives as the answer. In conditionals such as "come here or you are grounded", the point of the threat is that only one disjunct can be true: either "you come and you are not grounded" or "you don't come and you are grounded".

Repairs often received an exclusive (XOR) or a second-disjunct-true (NAB) interpretation. This is expected given that in repairs the speaker intends to say that the first disjunct is incorrect or inaccurate. Unconditionals and definitions/examples always had a conjunctive (AND) interpretation. Again, this is to be expected. In such cases the speaker intends to communicate that all options apply. If the mother says that "cats are animals like lions or tigers", she intends to say that both lions and tigers are cats, and not one or the other. Interestingly, in some cases, *or* is replaceable by *and*: "cats are animals like lions and tigers". In unconditionals, the speaker communicates that in both alternatives, a certain

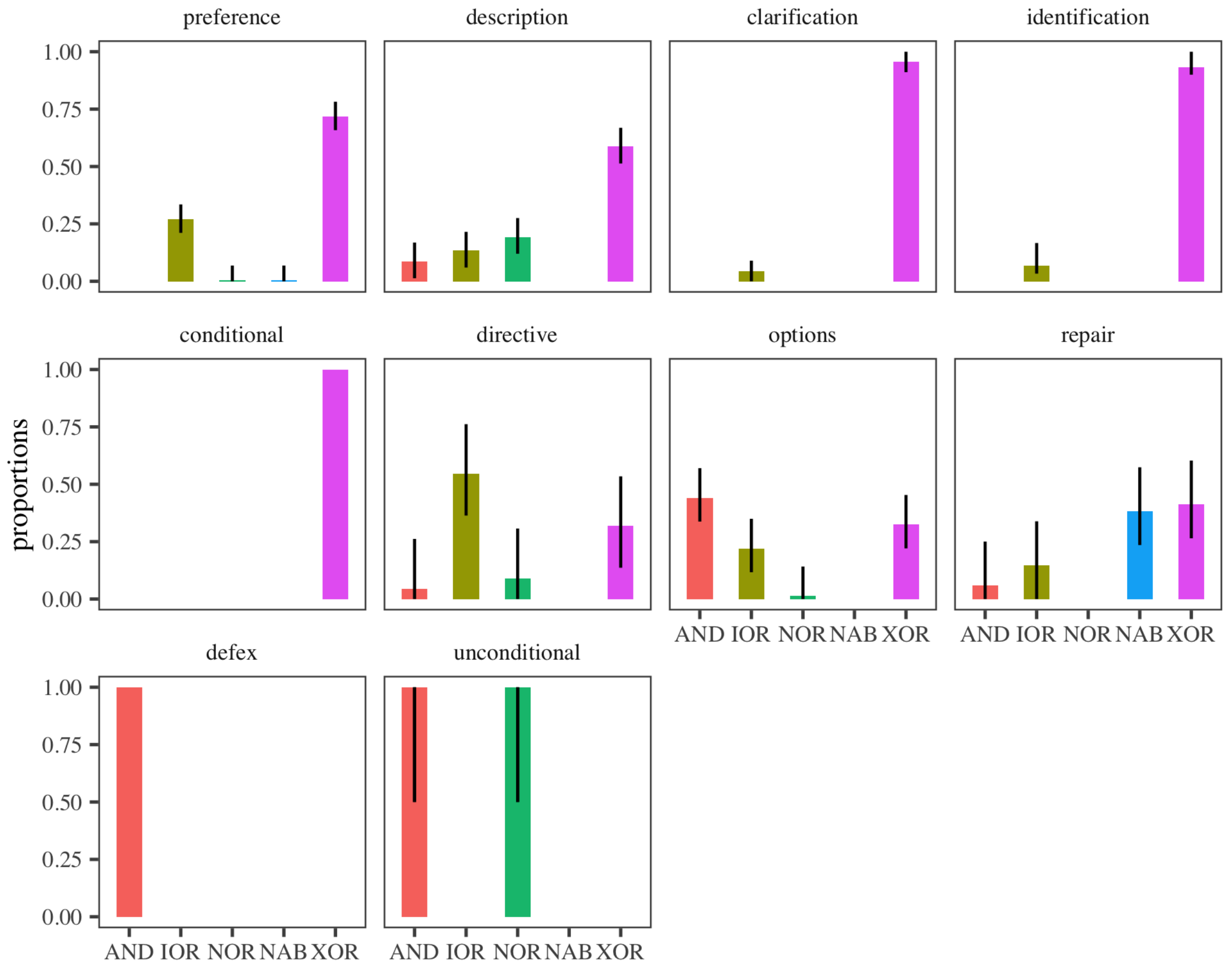

Figure 14. Interpretations of disjunction in different communicative functions.

proposition holds. For example, if the mother says "ready or not, here I come!", she communicates that "I come" is true in both cases where "you are ready" and "you are not ready".

Options were often interpreted either as conjunctive (AND) or inclusive (IOR). The category "options" contained examples of free-choice inferences such as "you could drink orange juice or apple juice". This study found free-choice examples to be more common in child-directed speech than the current literature on the acquisition of disjunction assumes. Finally, directives received either an IOR or XOR interpretation. It is important to note here that the most common communicative function in the data were preferences and descriptions. Other communicative functions such as unconditionals or options were fairly rare. Despite their infrequent appearance, these constructions must be learned by children at some point, since almost all adults know how to interpret them.

Conclusion

This study focused on the interpretations that connectives *and* and *or* recieve in child-directed speech. It also investigated some candidate cues that can help children's learning of these interpretations. The study selected 1000 examples of *and* and *or* in child-directed speech, annotated for their truth-conditional interpretation, as well as six candidate cues: (1) Conceptual Consistency (2) Utterance Type; (3) Intonation; (4) Presence of negative or modal morphemes (5) Syntactic Level; and (6) Communicative Function. Like Morris (2008), this study found that the most common interpretations of *and* and *or* are conjunction (AND) and exclusive disjunction (XOR) respectively. Therefore, relying only on connective word forms, a learner should expect *and* to be a conjunction and *or* exclusive disjunction.

However, the study also found that the most likely interpretation of a disjunction depended on the cues that accompanied it in context. A disjunction was most likely exclusive if the alternatives were inconsistent (i.e. contradictory). A disjunction with consistent alternatives was either inclusive or exclusive if it appeared in a question. Within questions, a disjunction was most likely exclusive if its intonation was “rise-fall”, and inclusive if it was “rising”. Among declaratives and imperatives with “flat” intonations, a disjunction was interpreated most liklely as AND if there was a modal, and NOR or XOR if there was negation present in the utterance. Finallly, in the absence of all these cues, a disjunction was more likely to be non-exclusive (IOR + AND) than exclusive (XOR). These results suggest that a learner can potentially use these cues to predict the intended interpretation of a connective in utterance context. In the next study, we use a computational learning model to formalize this account.

Study 3: Computational Modeling Using Decision Trees

A decision tree is a classification model structured as a hierarchical tree with an initial node, called the root, that branches into more nodes until it reaches the leaves (Breiman, 2017). Each node represents a test on a feature, each branch represents an outcome of the test, and each leaf represents a classification label. Using a decision tree, observations can be classified or labeled based on a set of features. Decision trees have at least three advantages for modeling cue-based accounts of semantic acquisition. First, the features used in decision trees for classification can be the cues that help the acquisition and interpretation of a word or an utterance. Second, unlike many other machine learning techniques, decision trees result in models that are interpretable. Third, the order of decisions or features used for classification is determined based on information gain. Features that appear higher (earlier) in the tree are more informative and helpful for classification. Therefore, decision trees can help us understand which cues are more helpful for the acquisition and interpretation of words.

Decision tree learning is the construction of a decision tree from labeled training data. This section applies decision tree learning to the annotated data of Study 2 by constructing random forests (Breiman, 2001; Ho, 1995). In random forest classification, multiple decision trees are constructed on subsets of the data, and each tree predicts a classification. The ultimate outcome is a majority vote of each tree's classification. Since decision trees tend to overfit data, random forests control for overfitting by building more trees and averaging their results. **(Citation)** In the context of semantic acquisition, the random forest can represent hypothetical variability in the learners. The next section discusses the methods used in constrcting the random forests for interpreting the connectives *or* and *and*.

Methods

The random forest models were constructed using python's Sci-kit Learn package (Pedregosa et al., 2011). The annotated data had a feature array and a connective interpretation label for each connective use. Connective interpretations included exclusive (XOR), inclusive (IOR), conjunctive (AND), negative inclusive (NOR), and NPQ which states that only the second proposition is true. The features or cues used included all other annotation categories: intonation, consistency, utterance type, syntactic level, negation, modals, and communicative function. All models were trained with stratified 10-Fold cross-validation to reduce overfitting. Stratified cross-validation maintains the distribution of the initial data in the random sampling to build cross validated models. Maintaining the data distribution ensures a more realistic learning environment for the forests. Tree success was measured with F1-Score, harmonic average of precision and recall (Rijsbergen, 1979).

First a grid search was run on the hyperparamter space to establish the number of trees in each forest and the maximum tree depth allowable. The grid search creates a grid of all combinations of forest size and tree depth and then trains each forest from this grid on the data. The forests with the best F1-score and lowest size/depth are reported. **(Citation*)** The default number of trees for the forests was set to 20, with a max depth of eight and a minimum impurity decrease of 0. Impurity was measured with gini impurity, which states the odds that a random member of the subset would be mislabled if it were randomly labeled according to the distribution of labels in the subset. (Gini, 1912).

Decision trees were fit with high and low minimum-gini-decrease values. High minimum-gini-decrease results in a tree that does not use any features for branching. Such a tree represents the baseline or traditional approach to mapping that directly maps a word to its most likely interpretation. Low minimum-gini-decrease allows for a less conservative tree that uses multiple cues or features to predict the interpretation of a disjunction. Such a tree represents the cue-based context-sensitive account of word learning.

Results

We first present the results of the random forests in the binary classification task. The models were trained to classify exclusivity, e.i. whether an interpretation was exclusive or not. For visualization of trees, we selected the highest performing tree in the forest by testing each tree and selecting for highest F1 score. While the forests performance is not identical to the highest performing tree, the best tree gives an illustrative example of successful learning from data. Figure 15 shows the best performing decision tree with high minimum gini decrease. As expected, a learner that does not use any cues would interpret *or* as exclusive all the time. This is the baseline model. Figure 16 shows the best performing decision tree with low minimum gini decrease. The tree has learned to use intonation and consistency to classify disjunctions as exclusive or inclusive. As expected, if the intonation is rise-fall or the disjuncts are inconsistent, the interpretation is exclusive. Otherwise, the disjunction is classified as not exclusive.

Figure **??** shows the average F1 scores of the baseline and cue-based models in classifying exclusive examples as the number of training examples increases. The models perform similarly, but the cue-based model performs slightly better (no significant difference). The real difference between the baseline model and the cue-based model is in their performance on inclusive examples. Figure **??** shows the F1 score of the forests as a function of the training size in classifying inclusive examples. As expected, the baseline model performs very poorly while the cue-based model improves with more examples and performs significantly better than the baseline tree.

Next, we use decision tree learning in a ternary classification task. The model uses features to interpret a coordination with *and* and *or* as inclusive (IOR), exclusive (XOR), or

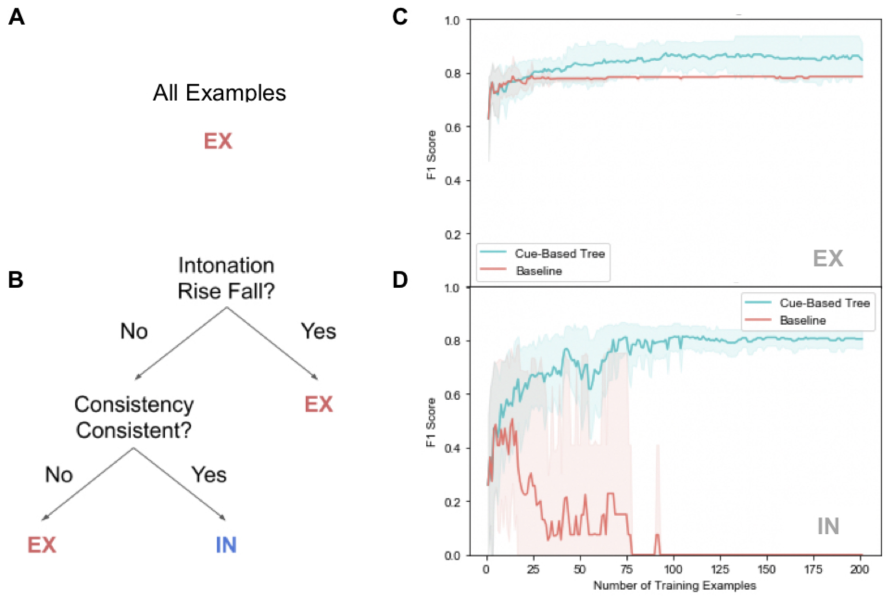

Figure 15. (A) The structure for the baseline (highest gini threshold, 0.2) decision tree trained on examples with exclusive (EX) and non-exclusive (IN) interpretations. (B) The structure for the cue-based decision tree (low gini threshold of 0.01). The average F1 score with 95% confidence intervals as a function of the number of training examples in the baseline and cue-based model when treating as positive (C) EX and (D) IN respectively.

conjunctive (AND). Figure **??** shows the baseline decision tree with high minimum gini decrease, which only uses the presence of the words *or*/*and* to interpret conjunction and disjunction. As expected, the tree interprets a coordination with *and* as a conjunction and one with *or* as exclusive disjunction. Figure **??** shows the cue-based decision tree with low minimum gini decrease. In addition to the presence of *and* and *or*, the tree uses intonation, consistency, communicative function, and utterance type to distinguish exclusive, inclusive, and conjunctive uses of disjunction. In short, a disjunction that is rise-fall, inconsistent, or has a conditional communicative function is classified as exclusive. Otherwise the disjunction is classified as inclusive. The tree also finds conjunctive interpretations of disjunction more likely in declarative sentences than interrogatives.

Figure **??** shows the average F1 score of the conjunctive interpretations (AND) for the baseline and the cue-based models. Since the vast majority of the conjunctive interpretations are predicted by the presence of the word *and*, the baseline and cue-based models show similar performances. Setting aside conjunction examples, Figure **??** shows the average F1 score of the AND interpretation of disjunction only. Here we see that the cue-based model performs better than the default model in guessing conjunctive interpretations of disjunction. The informal analysis of the trees suggest that the model does this by using the “speech act” cue. Figure **??** shows the average F1-score of the exclusive interpretations (XOR) for the baseline and the cue-based models. The cue-based model does slightly better than the baseline model. As before, the most important improvement comes in identifying inclusive examples. Figure **??** shows the average F1-score of the inclusive interpretations (IOR) for both baseline and cue-based models. The baseline model performs very poorly while the cue-based model is capable of classifying inclusive examples as well.

Finally, we look at decision trees trained on the annotation data to predict all the interpretation classes for disjunction: AND, XOR, IOR, NOR, and NPQ. Figure **??** shows the baseline model that only uses the words *and* and *or* to classify. As expected, *and* receives a conjunctive interpretation (AND) and *or* receives an exclusive interpretation (XOR). Figure **??** shows the best example tree of the cue-based model. The leaves of the tree show that it recognizes exclusive, inclusive, conjunctive, and even negative inclusive (NOR) interpretations of disjunction. How does the tree achieve that? Like the baseline model, the tree first asks about the connective used: *and* vs. *or*. Then like the previous models, it asks about intonation and consistency. If the intonation is rise-fall, or the disjuncts are inconsistent, the interpretation is exclusive. Then it asks whether the sentence is an interrogative or a declarative. If interrogative, it guesses an inclusive interpretation. This basically covers questions with a rising intonation. Then the tree picks declarative examples that have conditional speech act (e.g. “give me the toy or you’re grounded”) and labels them as exclusive. Finally, if negation is present in the sentence, the tree labels the disjunction as NOR.

Figures **??**, **??**, and **??** show the average F1-scores for the conjunctive (AND), exclusive (XOR), and inclusive (IOR) interpretations as a function of training size. The results are similar to what wereported before with the ternary classification. While the cue-based model generally performs better than the baseline model, it shows substantial improvement in classifying inclusive cases. Figure **??** shows the average F1-score for the negative inclusive interpretation as a function of training size. Compared to the baseline model, the cue-based model shows a substantially better performance in classifying negative sentences. The success of the model in classifying negative inclusive examples (NOR) suggests that the cue-based model offers a promising approach for capturing the scope relation of operators such as negation and disjunction. Here, the model learns that when negation and disjunction are present, the sentence receives a negative inclusive (NOR) interpretation. In other words, the model has learned the narrow-scope interpretation of negation and disjunction from the input data. In a language where negation and disjunction receive an XOR interpretation (not A or not B), the cue-based model can learn the wide-scope interpretation of disjunction.

Finally, Figure **??** shows the average F1 score for the class NPQ. This interpretation suggested that the first disjunct is false but the second true. It was seen in examples of repair most often and the most likely cue to it was also the communicative function or speech act of repair. The results show that even though there were improvements in the cue-based model, they were not stable as shown by the large confidence intervals. It is possible that with larger training samples, the cue-based model can reliably classify the NPQ interpretations as well.

Conclusion

In this study, we used the annotation data from Study 2 to train and compare two random forest models, representing two accounts for the acquisition of disjunction. The first account was a baseline (context-independent) account in which words are isolated and directly mapped to their most likely meanings, disregarding available contextual cues.

Random forest models with high minimum-gini-impurity-decrease represented this account. The second account was what we called the cue-based context-dependent mapping in which words are mapped to meanings using a set of cues available in the context. Random forest models with low minimum-gini-impurity-decrease represented this cue-based account. Comparison of the F1-Scores produced by models representing these two accounts showed that the cue-based models outperfromed the baseline models in every classification task. Most importantly, while the baseline models learned to always interpret a disjunction as exclusive, the cue-based models learned to interpret a disjunction as exclusive, inclusive, conjunctive, or negative inclusive (NOR), depending on the cues available in the input.

General Discussion

This paper presented three studies to support the claim that child-directed speech contains prosodic, conceptual, and linguistic cues that can aid the acquisition of linguistic disjunction. Study 1 presented the overall distribution of *or* and *and* in parents' and children's speech in CHILDES corpora. We found that children heared 1-2 examples of *or* in every thousand words parents produced. Children started producing *or* themselves between 18-30 months, and by 42 months they reached a rate of one *or* per thousand words. Study 2 showed that as Morris (2008) had found, the most common interpretation of *or* in child-directed speech is exclusive disjunction. However, we also found that exclusive interpretations were cued by prosodic and conceptual cues. In the absence of prosodic and conceptual cues to exclusivity, the interpretation of a disjunction was most likely non-exclusive. Finally, study 3 used decision-tree learning to show that a hypothetical learner can use prosodic, conceptual, and linguistic cues to predict the interpretation of a disjunction. It is importnat to note that while this study has shown the **potential utility** of conceptual, prosodic, and linguistic cues present in child-directed speech for the acquisition of disjunction, it has not actually established that children learning disjunction are sensitive to these cues, or that use them for learning. It is important for future experimental research to follow up and show that these cues are actually used by langauge learners in their acquisition of disjunction.

In what follows in this section, we place our findings and main claim in the bigger context of word learning. As we mentioned at the beginning, theories of word learning have been heavily influenced by Quine's Gavagai thought experiment; however, his specific views on word learning are not as widely discussed. Quine (1960) pinpointed several dimensions of word learning where theories can differ in their approach. First, Quine considered form-meaning mapping to occur at all levels of linguistic structure: words, phrases, and sentences. In other words, the learner could map any recognizable chunk or construction to candidate meanings, and store it in memory. This is in contrast to a view where only morphemes (smallest units) are mapped and stored, while larger units are derived compositionally. We call this dimension the **mapping unit**. Second, Quine considered three different ways of mapping words to meanings: isolated, context-dependent, and described. "Isolated mapping" refers to the case of hearing a word such as "chair", "red", or "run" and mapping it to the percieved object, property, or action isolated from its linguistic or communicative context. It is the classic Gavagai example. On the other hand, "context-dependent mapping" is learning a word "contextually, or by abstraction, as a fragment of sentences learned as wholes". He suggested that "prepositions, conjunctions, and many other words, are bound to have been learned only contextually; we get on to using them by analogy with the ways in which they have been seen to turn up in past sentences". According to Quine, learning such words requires attention to the linguistic context of use. "Description mapping" is the extreme case in which a word is defined using other words only, similar to a dictionary entry. Quine points out that the meaning of a word such as "molecule" is mapped to a linguistic description (i.e. definition). We can call these three ways **mapping types**.

We add three more dimensions to the ones dicussed by Quine. Theories of form-meanign mapping may or may not rely on cues, and when they do, they may consider different roles for the cues. We call this dimension **cue status**. For example, substantitive nominals (to use Quine's terminology) are hypothesized to benefit from social cues such as eye gaze and pointing. Verbs are hypothesized to benefit from syntactic cues, and in this paper we argued for conceptual and linguistic cues for the acquisition of disjunction. However, the role of these cues are not the same. For example, in mapping nominals, eye gaze and pointing act as cues that enhance the odds of a particular mapping. However, in the research presented here, conceptual and linguistic cues help to partition the input and better specify the context of use.

Theories of form-meaning mapping also differ in their coneptual-representational **primitives**, e.i. the units that linguistic forms are mapped to. For exmaple, in Morris (2008)'s account, the semantic space for connectives included temporal and causal conjunction, as well as exclusive disjunction. The nativist account, on the other hand, does not assume these primitives (Crain, 2012). The choice of primitives has a crucial role in word learning. For example, the nativist account resolves the puzzle of learning disjunction by positing primitives that exclude exclusive disjunction. This way, a word like *or* can only be mapped to inclusive disjunction and the exclusive interpretations are byproducts of pragmatic computations. In the account presented in this paper, we constrained our semantic primitives to the 16 logical (truth-functional) connectives, and argued that conceptual and linguistic cues can help a learner learn both exclusive and inclusive disjunction using the input data. Is there any reason to believe that learners can constrain the hypothesis space to connective meanings while mapping a word like *or*? We believe so. Connectives have a very specific syntactic distribution, and in our data, majority of *and* and *or* examples were used to connect clauses. However, we leave the precise mechanism of selecting specific functional domains in the hypothesis space such as connective meanings using syntactic information for future work.

Finally, theories of form-meaning mapping may differ on their assumptions on **conceptual continuity**. Constructivist accounts emphasize conceptual development and construction of thought from a non-adult-like early stage. Many nativist accounts, on the other hand, assume that concepts used in early form-meaning mapping are similar to those used by adults. An important step in providing evidence for such conceptual continuity has been to show that children's early interpretations correspond to adult semantics in other languages. In its current version, our account of disjunction assumes continuity. However, it is possible to develop a version in which the primitive concepts in this model are developed from other social or perceptual primitives. A possible social primitive for the concept of disjunction is "choice between two or more alternatives" (Braine & Rumain, 1981). However, a concrete proposal with specific predictions for the developmental stages has not been developed yet.

To summarize, the account presented in this paper for the acquisition of disjunciton is cue-based and context-dependent. It assumes that the learner has the 16 binary logical connective concepts available for mapping to linguistic forms. For its mapping units, it goes beyond mapping the word *or* in isolation to a hypothesized meaning, and stores information about the conceptual and linguistic context of the word as well. However, it does not record all the information content of an utterance either. Finally, in its current format, our account assumes conceptual continuity. Most importantly, our study shows that such an account resolves the puzzle of learning disjunction and obviates the need for a more constrained hypothesis space that excludes exclusive disjunction.

References

Appendix

Table 2

Information on the participants in the Providence Corpus. Ethan was diagnosed with Asperger's syndrome and therefore was excluded from this study.

Name	Age Range	Sessions
Alex	1;04.28-3;05.16	51
Ethan	0;11.04-2;11.01	50
Lily	1;01.02-4;00.02	80
Naima	0;11.27-3;10.10	88
Violet	1;02.00-3;11.24	51
William	1;04.12-3;04.18	44

Annotation Categories

Table 3

Annotation classes for connective interpretation

Class	Meaning	Examples
AND	Both propositions are true	*"I'm just gonna empty this and then I'll be out of the kitchen." – "I'll mix them together or I could mix it with carrot, too."*
IOR	One or both propositions are true	*"You should use a spoon or a fork." – "Ask a grownup for some juice or water or soy milk."*
XOR	Only one proposition is true	*"Is that a hyena? or a leopard?" – "We're gonna do things one way or the other."*

Class	Meaning	Examples
NOR	Neither proposition is true	*"I wouldn't say boo to one goose or three." – "She found she lacked talent for hiding in trees, for chirping like crickets, or humming like bees."*
IFF	Either both propositions are true or both are false	*"Put them [crayons] up here and you can get down. – Come over here and I'll show you."*
NAB	The first proposition is false, the second is true.	*"There's an Oatio here, or actually, there's a wheat here."*

Table 4

Definitions of the intonation types and their examples.

Intonation	Definitions	Examples
Flat	Intonation does not show any substantial rise at the end of the sentence.	*"I don't hear any meows or bow-wow-wows."*
Rise	There is a substantial intonation rise on each disjunct or generally on both.	*"Do you want some seaweed? or some wheat germ?"*
Rise-Fall	There is a substantial rise on the non-final disjunct(s), and a fall on the final disjunct.	*"Is that big Q or little q?" – "(are) You patting them, petting them, or slapping them?"*

Table 5

Definitions of the utterance types and their examples.

Utterance Types	Definitions	Examples
Declarative	A statement with a subject-verb-object word order and a flat intonation.	*"It looks a little bit like a drum stick or a mallet."*
Interrogative	A question with either subject-auxiliary inversion or a rising terminal intonation.	*"Is that a dog or a cat?"*
Imperative	A directive with an uninflected verb and no subject	*"Have a little more French toast or have some of your juice."*

Table 6

Definitions of the syntactic levels and their examples.

Syntactic Level	Definitions	Examples
Clausal	The coordinands are sentences, clauses, verb phrases, or verbs.	*"Does he lose his tail sometimes and Pooh helps him and puts it back on?"*
Sub-clausal	The coordinands are nouns, adjectives, noun phrases, determiner phrases, or prepositional phrases.	*"Hollies can be bushes or trees."*

Table 7

Definitions of consistency types and their examples.

Consistency	Definitions	Examples
Consistent	The coordinands can be true at the same time.	*"We could spell some things with a pen or draw some pictures."*
Inconsistent	The coordinands cannot be true at the same time.	*"Do you want to stay or go?"*

Table 8

Definitions of the communicative functions and their examples.

Function	Definitions	Examples
Descriptions	Describing what the world is like or asking about it. The primary goal is to inform the addressee about how things are.	"*It's not in the ditch or the drain pipe.*"
Identifications	Identifying the category membership or an attribute of an object. Speaker has uncertainty. A subtype of "Description".	"*Is that a ball or a balloon honey?*"
Definitions and Examples	Providing labels for a category or examples for it. Speaker is certain. Subtype of Description.	*"This is a cup or a mug." – "berries like blueberry or raspberry"*
Preferences	Asking what the addressee wants or would like or stating what the speaker wants or would like	*"Do you wanna play pizza or read the book?"*

Function	Definitions	Examples
Options	Either asking or listing what one can or is allowed to do. Giving permission, asking for permission, or describing the possibilities. Often the modal "can" is either present or can be inserted.	*"You could have wheat or rice."*
Directives	Directing the addressee to act or not act in a particular way. Common patterns include "let's do ...", "Why don't you do ...", or prohibitions such as "Don't ...". The difference with "options" is that the speaker expects the directive to be carried out by the addressee. There is no such expectation for "options".	*"let's go back and play with your ball or we'll read your book."*
Clarifications	Something is said or done as a communicative act but the speaker has uncertainty with respect to the form or the content.	*"You mean boba or bubble?"*
Repairs	Speaker correcting herself on something she said (self repair) or correcting the addressee (other repair). The second disjunct is what holds and is intended by the speaker. The speaker does not have uncertainty with respect to what actually holds.	*"There's an Oatio here, or actually, there's a wheat here."*

Function	Definitions	Examples
Conditionals	Explaining in the second coordinand, what would follow if the first coordinand is (or is not) followed. Subtype of Directive.	*"Put that out of your mouth, or I'm gonna put it away."* – *"Come over here and I'll show you."*
Unconditionals	Denying the dependence of something on a set of conditions. Typical format: "Whether X or Y, Z". Subtype of Descriptions.	*"Ready or not, here I come!"* (playing hide and seek)

Table 9

Definitions of answer types and their examples.

Type	Definitions	Examples
No Answer	The child provides no answer to the question.	Mother: *"Would you like to eat some applesauce or some carrots?"* Child: *"Guess what Max!"*
YN	The child responds with *yes* or *no*.	Father: *"Can I finish eating one or two more bites of my cereal?"* Child: *"No."*
AB	The child responds with one of the disjuncts (alternatives).	Mother: *"Is she a baby elephant or is she a toddler elephant?"* Child: *"It's a baby. She has a tail."*

Inter-annotator agreement

Figure 17 shows the percentage agreement and the kappa values for each annotation category over the 8 iterations.

Agreement in the following three categories showed substantial improvement after better and more precise definitions and annotation criteria were developed: connective interpretation, intonation, and communicative function. First, connective interpretation showed major improvements after annotators developed more precise criteria for selecting the propositions under discussion and separately wrote down the two propositions connected by the connective word. For example, if the original utterance was “do you want milk or juice?”, the annotators wrote “you want milk, you want juice” as the two propositions under discussion. This exercise clarified the exact propositions under discussion and sharpened annotator intuitions with respect to the connective interpretation that is communicated by the utterance. Second, annotators improved agreement on intonation by reconstructing an utterance’s intonation for all three intonation categories. For example, the annotator would examine the same sentence “do you want coffee or tea?” with a rise-fall, a rise, and a flat intonation. Then the annotator would listen to the actual utterance and see which one most resembled the actual utterance. This method helped annotators judge the intonation of an utterance more accurately. Finally, agreement on communicative functions improved as the definitions were made more precise. For example, the definition of “directives” in Table 8 explicitly mentions the difference between “directives” and “options”. Clarifying the definitions of communicative functions helped improve annotator agreement.

Inter-annotator reliability for conjunction was calculated in the same way. Two different annotators coded 300 utterances of *and*. Inter-annotator reliability was calculated over 10 iterations of 30 examples. Figure 18 shows the percentage agreement between the annotators as well as the kappa values for each iteration. Despite high percentage agreement between annotators, the kappa values did not pass the set threshold of 0.7 in three consecutive iterations. This paradoxical result is mainly due to a property of kappa. An imbalance in the prevalence of annotation categories can drastically lower its value. When one category is extremely common with high agreement while other categories are rare, kappa will be low (Cicchetti & Feinstein, 1990; Feinstein & Cicchetti, 1990). In almost all annotated categories for conjunction, there was one class that was extremely prevalent. In such cases, it is more informative to look at the class specific agreement for the prevalent category than the overall agreement measured by Kappa (Cicchetti & Feinstein, 1990; Feinstein & Cicchetti, 1990).

Table 10 lists the dominant classes as well as their prevalence, the values of class specific agreement index, and category agreement index (Kappa). Class specific agreement index is defined as $2n_{ii}/n_{i.} + n_{.i}$, where i represents the class's row/column number in the category's confusion matrix, n the number of annotations in a cell, and the dot ranges over all the row/column numbers (Fleiss, Levin, & Paik, 2013, p. 600; Ubersax, 2009). The class specific agreement indices are high for all the most prevalent classes showing that the annotators had very high agreement on these class, even though the general agreement index (Kappa) was often low. The most extreme case is the category "consistency" where almost all instances were annotated as "consistent" with perfect class specific agreement but low overall Kappa. In the case of utterance type and syntactic level where the distribution of instances across classes was more even, the general index of agreement Kappa is also high. In general, examples of conjunction showed little variability across annotation categories and mostly fell into one class within each category. Annotators had high agreement for these dominant classes.

Baldwin, D. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language*, *20*(2), 395–418.

Braine, M. D., & Rumain, B. (1981). Development of comprehension of "or": Evidence for a sequence of competencies. *Journal of Experimental Child Psychology*, *31*(1), 46–70.

Table 10

Most prevalent annotation class in each annotation category with the values of class agreement indeces and category agreement indeces (Kappa).

Annotation Category	Class	Prevalence	Class Agreement Index	Kappa
intonation	flat	0.86	0.89	0.24
interpretation	AND	0.96	0.98	0.39
answer	NA	0.84	0.94	0.67
utterance_type	declarative	0.76	0.94	0.70
communicative_function	description	0.77	0.90	0.59
syntactic_level	clausal	0.67	0.91	0.70
consistency	consistent	0.99	1.00	0.50

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Breiman, L. (2017). *Classification and regression trees.* London: Routledge.

Brown, R. W. (1957). Linguistic determinism and the part of speech. *The Journal of Abnormal and Social Psychology*, *55*(1), 1.

Chierchia, G., Crain, S., Guasti, M. T., Gualmini, A., & Meroni, L. (2001). The acquisition of disjunction: Evidence for a grammatical view of scalar implicatures. In *Proceedings of the 25th Boston University conference on language development* (pp. 157–168). Somerville, MA: Cascadilla Press.

Chierchia, G., Guasti, M. T., Gualmini, A., Meroni, L., Crain, S., & Foppolo, F. (2004). Semantic and pragmatic competence in children's and adults' comprehension of or. In I. Noveck & D. Sperber (Eds.), *Experimental pragmatics* (pp. 283–300). Basingstoke: Palgrave Macmillan.

Cicchetti, D. V., & Feinstein, A. R. (1990). High agreement but low kappa: II. Resolving the paradoxes. *Journal of Clinical Epidemiology*, *43*(6), 551–558.

Clark, E. V. (1993). *The lexicon in acquisition*. Cambridge University Press.

Clark, E. V. (2009). *First language acquisition* (2nd ed.). Cambridge: Cambridge University Press.

Crain, S. (2012). *The emergence of meaning*. Cambridge: Cambridge University Press.

Crain, S., Gualmini, A., & Meroni, L. (2000). The acquisition of logical words. *LOGOS and Language*, *1*, 49–59.

Crain, S., & Khlentzos, D. (2008). Is logic innate? *Biolinguistics*, *2*(1), 024–056.

Crain, S., & Khlentzos, D. (2010). The logic instinct. *Mind & Language*, *25*(1), 30–65.

Demuth, K., Culbertson, J., & Alter, J. (2006). Word-minimality, epenthesis and coda licensing in the early acquisition of English. *Language and Speech*, *49*(2), 137–173.

Feinstein, A. R., & Cicchetti, D. V. (1990). High agreement but low kappa: I. The problems of two paradoxes. *Journal of Clinical Epidemiology*, *43*(6), 543–549.

Fleiss, J. L., Levin, B., & Paik, M. C. (2013). *Statistical methods for rates and proportions*. New York: John Wiley & Sons.

Geurts, B. (2006). Exclusive disjunction without implicatures. *Ms., University of Nijmegen*.

Gini, C. (1912). Variabilità e mutabilità. *Reprinted in Memorie Di Metodologica Statistica (Ed. Pizetti E, Salvemini, T). Rome: Libreria Eredi Virgilio Veschi*.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, *1*(1), 3–55.

Goodman, J. C., Dale, P. S., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language*, *35*(3), 515–531.

Grice, H. P. (1989). *Studies in the way of words.* Cambridge, MA: Harvard University Press.

Gualmini, A., & Crain, S. (2002). Why no child or adult must learn de Morgan's laws. In *Proceedings of the Boston University conference on language development.* Somerville, MA: Cascadilla Press.

Gualmini, A., Crain, S., & Meroni, L. (2000). Acqisition of disjunction in conditional sentences. In *Proceedings of the boston university conference on language development.*

Haspelmath, M. (2007). Coordination. In T. Shopen (Ed.), *Language typology and linguistic description,* Cambridge: Cambridge University Press.

Ho, T. K. (1995). Random decision forests. In *Proceedings of the third international conference on document analysis and recognition* (Vol. 1, pp. 278–282). Washington, DC, USA: IEEE Computer Society.

Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R., Brown, E., Chung, H. L., ... Rocroi, C. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, *65*(3), 1–123.

Johansson, B. S., & Sjolin, B. (1975). Preschool children's understanding of the coordinators "and" and "or". *Journal of Experimental Child Psychology*, *19*(2), 233–240.

Kamp, H. (1973). Free choice permission. In *Proceedings of the Aristotelian society* (Vol. 74, pp. 57–74).

Levy, E., & Nelson, K. (1994). Words in discourse: A dialectical approach to the acquisition of meaning and use. *Journal of Child Language*, *21*(02), 367–389.

MacWhinney, B. (2000). *The CHILDES project: The database* (Vol. 2). Mahwah, NJ: Erlbaum.

Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, *14* (1), 57–77.

Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology*, *16* (1), 1–27.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20* (2), 121–157.

Morris, B. J. (2008). Logically speaking: Evidence for item-based acquisition of the connectives "and" and "or". *Journal of Cognition and Development*, *9* (1), 67–88.

Neisser, U., & Weene, P. (1962). Hierarchies in concept attainment. *Journal of Experimental Psychology*, *64* (6), 640.

Notley, A., Thornton, R., & Crain, S. (2012a). English-speaking children's interpretation of disjunction in the scope of "not every". *Biolinguistics*, *6* (1), 32–69.

Notley, A., Zhou, P., Jensen, B., & Crain, S. (2012b). Children's interpretation of disjunction in the scope of "before": A comparison of English and Mandarin. *Journal of Child Language*, *39* (03), 482–522.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … others. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, *12* (Oct), 2825–2830.

Pruitt, K., & Roelofsen, F. (2013). The interpretation of prosody in disjunctive questions. *Linguistic Inquiry*, *44* (4), 632–650.

Quine, W. V. O. (1960). *Word and object.* Cambridge, MA: MIT press.

Rijsbergen, C. J. V. (1979). *Information retrieval* (2nd ed.). Newton, MA, USA: Butterworth-Heinemann.

Sanchez, A., Meylan, S., Braginsky, M., MacDonald, K., Yurovsky, D., & Frank, M. C. (2018). Childes-db: A flexible and reproducible interface to the child language data exchange system. PsyArXiv. Retrieved from psyarxiv.com/93mwx

Siskind, J. M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, *61*(1-2), 39–91.

Sison, C. P., & Glaz, J. (1995). Simultaneous confidence intervals and sample size determination for multinomial proportions. *Journal of the American Statistical Association*, *90*(429), 366–369.

Smith, K., Smith, A. D., & Blythe, R. A. (2011). Cross-situational learning: An experimental study of word-learning mechanisms. *Cognitive Science*, *35*(3), 480–498.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition.* Harvard University Press.

Ubersax, J. (2009). Retrieved from http://www.john-uebersax.com/stat/raw.htm

Von Wright, G. H. (1968). An essay in deontic logic and the general theory of action.

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, *18*(5), 414–420.

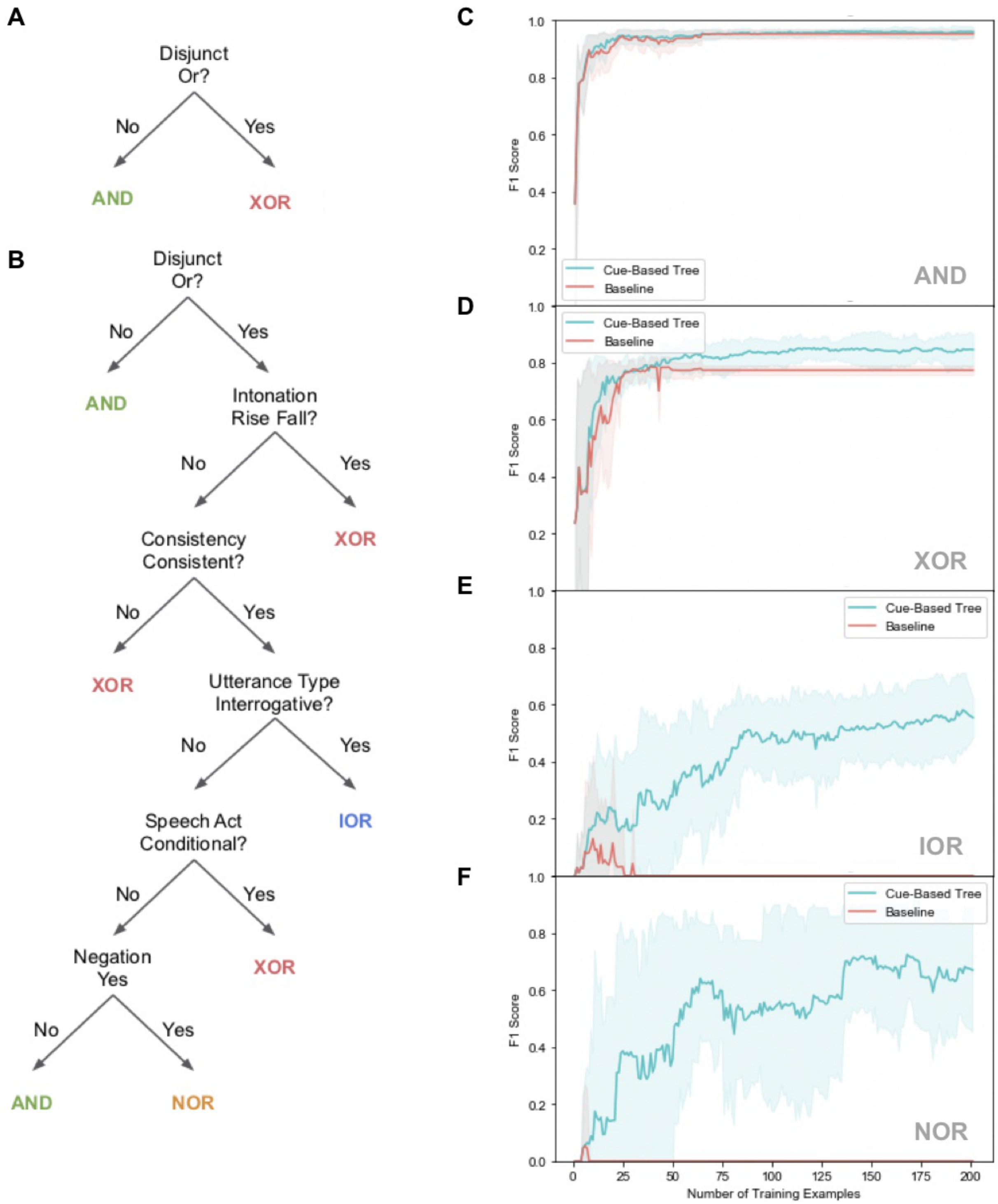

Figure 16. (A) The structure for the baseline (highest gini threshold, 0.2) decision tree trained on examples with XOR, IOR, AND, and NOR interpretations. (B) The structure for the cue-based decision tree (low gini threshold of 0.01). The average F1 score with 95% confidence intervals as a function of the number of training examples in the baseline and cue-based model when treating as positive (C) AND, (D) XOR, (E) IOR, (F) NOR respectively.

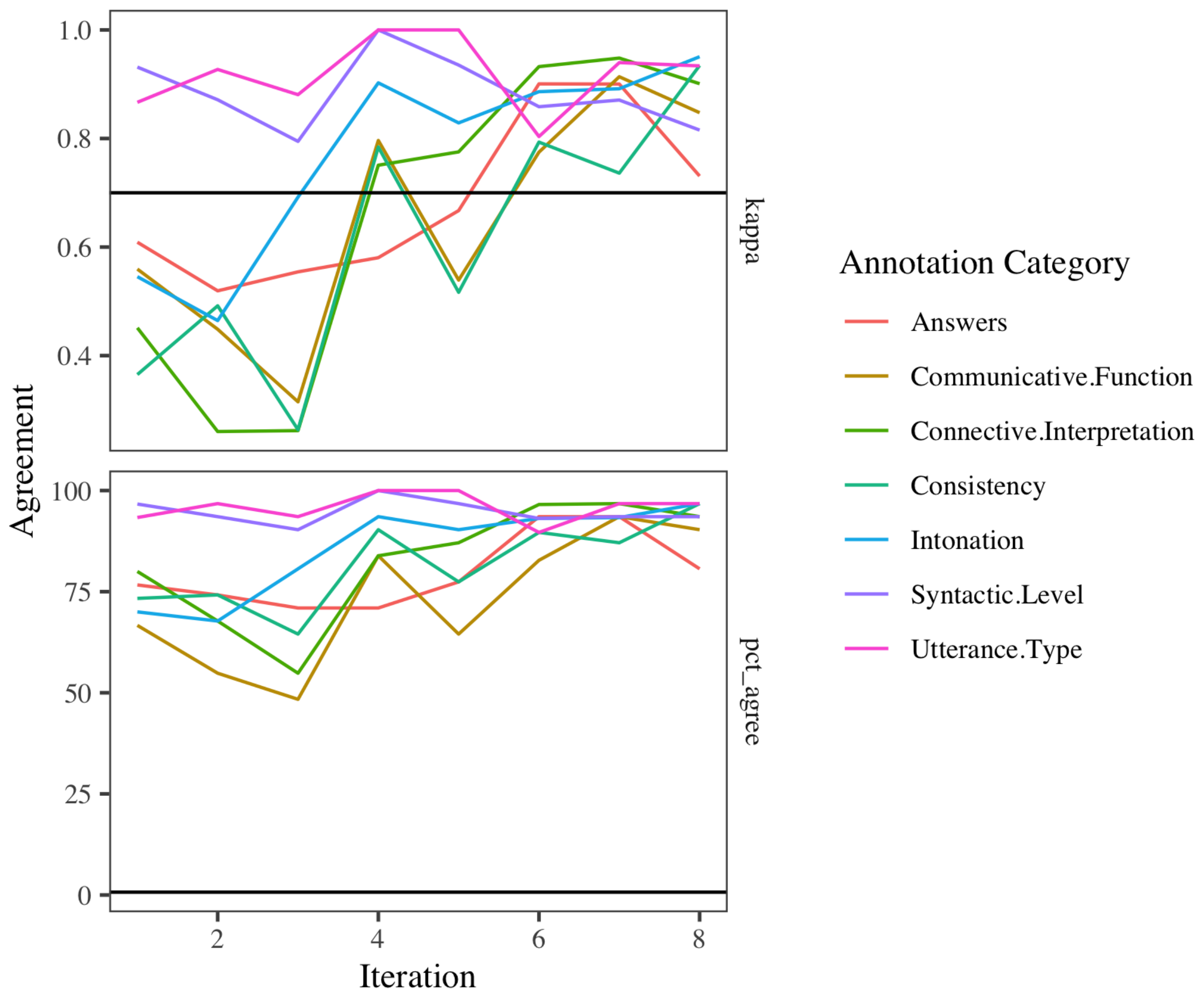

Figure 17. Inter-annotator agreement for disjunction examples.

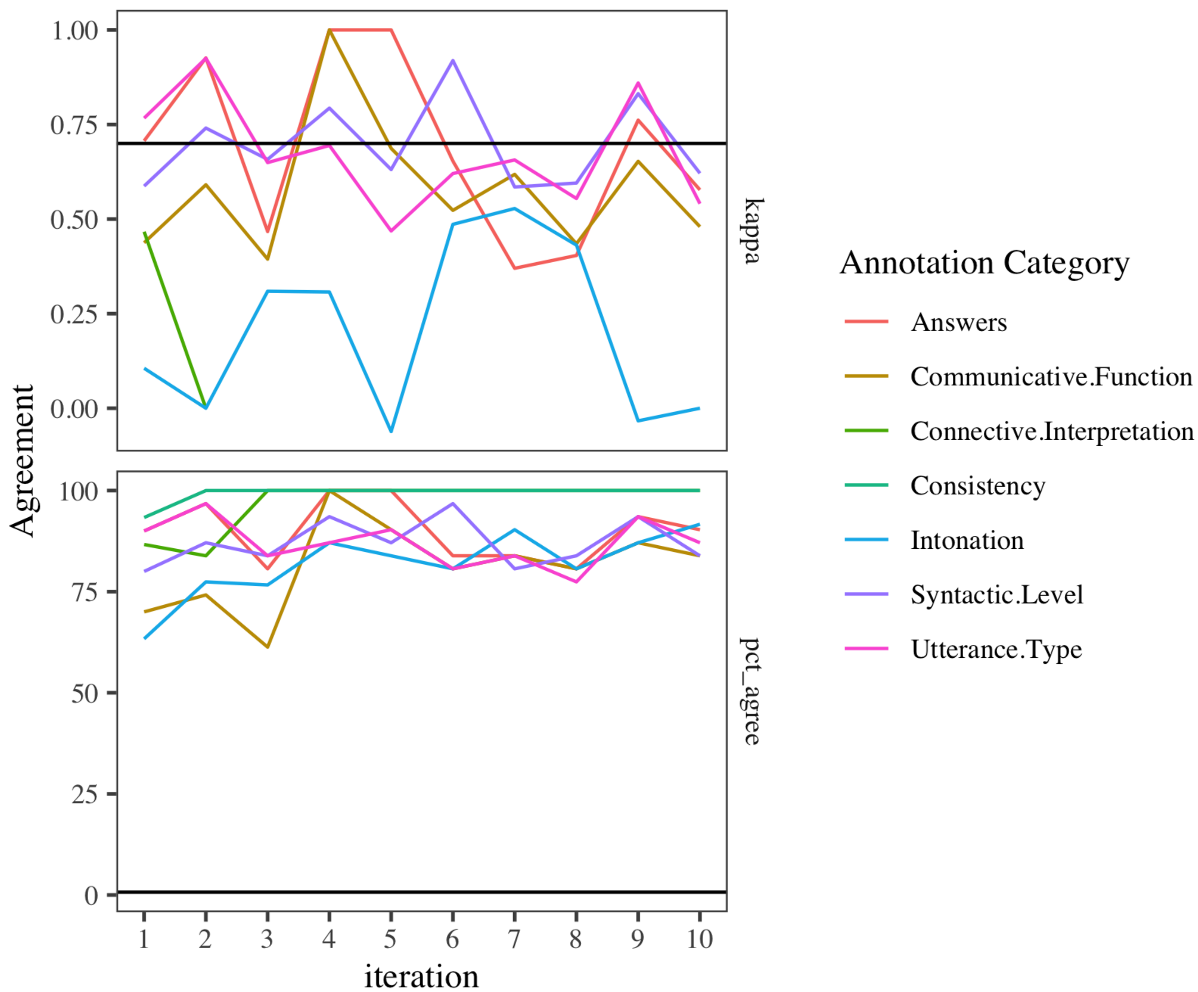

Figure 18. Inter-annotator agreement for conjunction examples.